

THÈSE

pour obtenir le grade de docteur délivré par

École Doctorale pluridisciplinaire de l'UAG

"Santé, Environnement et Sociétés dans les Amériques"

Spécialité doctorale "Mathématiques appliquées"

présentée et soutenue publiquement par

Clara LANDRY

le 18 Mai 2015

**Modélisation des dynamiques de maladies foliaires de
cultures pérennes tropicales à différentes échelles spatiales :
cas de la cercosporiose noire du bananier**

Directeur de thèse : **Jean VAILLANT**

Co-encadrants de thèse : **Catherine ABADIE** et **François BONNOT**

Jury

Dr. David MAKOWSKI,	INRA AgroParisTech, Thiverval-Grignon	Rapporteur
Pr. émérite Michel LANGLAIS,	Institut de Mathématiques, Université de Bordeaux	Rapporteur
Dr. Agnès CALONNEC,	INRA, Bordeaux	Examineur
Pr. émérite Gilbert SAPORTA,	CNAM, Paris	Président du jury
Pr. Jean VAILLANT,	Laboratoire de Mathématiques, UA Guadeloupe	Directeur de thèse
Dr. Catherine ABADIE,	CIRAD, Guadeloupe	Co-encadrant de thèse
M. François BONNOT,	CIRAD, Montpellier	Co-encadrant de thèse

Remerciements

Cette thèse est le fruit d'un peu plus de trois années de travail et son aboutissement doit beaucoup à l'aide et au soutien de nombreuses personnes.

Je voudrais tout d'abord remercier mon directeur de thèse, Mr Jean Vaillant, d'avoir accepté d'encadrer ce travail et de m'avoir donné de précieux conseils méthodologiques.

Je voudrais ensuite remercier mes encadrants, Catherine Abadie et François Bonnot pour leur présence et leur soutien pendant toutes ces années:

Merci à Catherine de m'avoir proposé cette thèse, de m'avoir initiée à la phytopathologie. Merci pour ta présence du début aux dernières minutes de la rédaction de ce manuscrit, merci de ta patience et de ta disponibilité, de ton soutien dans les moments les plus intenses car sans toi rien n'aurait été possible et cette thèse ne se serait sans doute pas terminée !

Merci à François, d'avoir accepté d'assurer l'encadrement de cette thèse, de m'avoir transmis une partie de ses connaissances en modélisation, en informatique (merci de ta patience dans mon apprentissage du C++, c'était pas gagné!), de m'avoir accueillie pendant mes séjours à Montpellier, et merci pour tes relectures attentives de ce manuscrit à des heures tardives.

Merci aux membres du jury d'avoir accepté de juger mes travaux; à M. Makowski et M. Langlais pour leur rôle de rapporteurs, et à Mme Agnès Calonnec et M. Gilbert Saporta en tant qu'examinateurs.

Je remercie aussi les membres de mon comité de pilotage pour leurs conseils : Christian Lannou, Jean Carlier, Virginie Ravigné, Christophe Abraham, Meili Baragatti et Yves Dumont.

Je remercie aussi Françoise Carreel et Marie-Françoise Zapater pour leurs formations au logiciel "Visilog" et Samuel Ralambondrainy pour sa patience lors de nos demandes de modification du logiciel ; un grand merci aussi à Raphaël Pasquier de m'avoir aidé à apprivoiser la "bête" Wahoo !

Je tiens à remercier Emmanuel Salomon, Yolande Chillin-Charles, Térésa-Reina Martinez et Aurore Cavalier pour l'aide apportée dans les expérimentations en République Dominicaine et Sophie Veran pour l'appui sur la modélisation de l'invasion.

Un grand merci aussi à Pierre-Yves pour son soutien et ses gâteaux de champion, à Aurore, Olivier, Eléa, Maryne, Henry, Chloé, Esther, Pauline, Claude, Thierry, Cyril, Charlène, Guillaume, Jean-Phi, Béranger, Benoît, Camille, Fred, Seb et Florence pour le quotidien.

Merci aux amis, toujours là pour me soutenir et pour organiser des soirées auxquelles je ne pouvais pas assister ! : Erwann, Marion, Rémi, Loic, Magaly, Myriam, Khalil, Fred, Seb, Michel, Cam, Flo, Liliane, Thomas. Merci aussi aux amis de longue date, Loulou, Fab, Marion, Jo, Noémie, Filou, Alek, Gui, Nico, Martin, Laura, Camille, Céline, Pauline.

Merci à ma famille, à ma mère pour ses encouragements, ses conseils et petites phrases à méditer qui m'ont beaucoup aidée, merci à mon père de m'avoir toujours soutenue et de s'être occupé de moi, même à distance. Merci aussi à Rémi, Christelle, Dany, Nicole, à ma grand-mère, à Paul et Martine, pour leurs encouragements. Merci à Carole, Seb, Nico et Lilou pour les petits séjours Martiniquais bien relaxants. Merci à Malie, André, Arthur et Agathe pour leur soutien, et merci à mon frère pour ses passages éclairs revivifiants.

Sans oublier tous ceux que je n'ai pas cités mais qui m'auront accompagnée durant ces années.

Enfin, cette thèse m'aura permis de rencontrer une personne exceptionnelle, Thomas, mon amour, merci pour ta présence, ton soutien sans failles et tes bons petits plats, je sais que le temps a été long pour toi, mais maintenant... la vie est à nous...

A ma famille,
A Thomas,

Résumé

Cette thèse concerne la modélisation des dynamiques de maladies foliaires de cultures pérennes tropicales à différentes échelles spatiales. Cette approche de modélisation est appliquée à la cercosporiose noire du bananier. Il s'agit d'explorer et de déterminer les paramètres environnementaux et de résistance de l'hôte ayant une influence significative sur la dynamique spatio-temporelle de la maladie et d'apporter des éléments associés au contrôle de la cercosporiose noire.

Deux modèles ont été développés dans le cadre de cette thèse. La dynamique épidémiologique au niveau de la plante est décrite par un modèle mécaniste décomposé en un modèle de croissance de la plante et un module épidémiologique décrivant le cycle épidémique du pathogène. L'architecture du bananier est prise en compte par le biais de compartiments foliaires positionnés dans l'espace. Le modèle a été validé par un jeu de données indépendant. Les expérimentations numériques et l'analyse de sensibilité du modèle réalisées par les méthodes de Morris et de e-FAST ont permis de mieux comprendre le fonctionnement épidémique de cette maladie et d'identifier les paramètres influant le plus la dynamique épidémique en particulier la vitesse d'extension des lésions, la durée d'incubation et l'efficacité d'infection. L'approche bayésienne a permis de prendre en compte l'information a priori disponible pour ces trois paramètres sur lesquels a porté l'inférence statistique. L'analyse de sensibilité du modèle a également permis d'identifier l'influence de deux paramètres liés à la croissance de la plante : le nombre de feuilles sur une plante et le rythme d'émission foliaire.

Un modèle de dynamique spatio-temporelle de la cercosporiose noire a été développé à l'échelle d'un territoire à partir d'enquêtes effectuées en Martinique pendant la période d'invasion de cette maladie de septembre 2010 à mai 2012. Les données récoltées étant censurées, l'inférence des paramètres du modèle a été réalisée dans un cadre bayésien en utilisant un algorithme d'augmentation de données. Le modèle et l'inférence développés permettent de reconstruire la dynamique spatio-temporelle de l'invasion et de prédire la fin d'invasion sur le territoire.

Les deux modèles de dynamiques spatio-temporelles développés à deux échelles spatiales différentes ont permis d'acquérir des informations importantes pour construire des outils de conception de méthode de contrôle de la cercosporiose noire des bananiers.

Table des matières

1	Introduction	1
1.1	Épidémies végétales et invasions biologiques	3
1.1.1	Définitions	3
1.1.2	Facteurs influant sur les dynamiques épidémiques	4
1.1.3	Modélisation spatio-temporelle et importance de l'échelle spatiale	5
1.2	Méthodes statistiques utilisées	5
1.2.1	Modélisation à l'échelle d'une plante	6
1.2.2	Modélisation à l'échelle d'un paysage	7
1.2.3	Méthodes d'estimation des paramètres	8
1.2.4	Analyse de sensibilité de modèles	11
1.3	Notre cas d'étude : la cercosporiose noire du bananier	13
1.3.1	Le bananier	13
1.3.2	La cercosporiose noire	14
1.3.3	Le pathogène <i>Mycosphaerella fijiensis</i>	15
1.3.4	Épidémiologie	16
1.3.5	Origine et distribution	18
1.3.6	Méthodes de lutte	18
1.3.7	Modélisation de la cercosporiose noire du bananier	19
1.4	Objectifs de la thèse	20
2	Modèle de simulation de la cercosporiose noire à l'échelle d'un bananier : le modèle Si-BaToka	23
2.1	Article en préparation	24
2.2	Simulation et évaluation du modèle	70
2.2.1	Description de la plante	71
2.2.2	Densité d'inoculum primaire	73
2.2.3	Dynamique des inoculum au cours des cycles de culture	76
2.2.4	Effet de l'inoculum primaire sur la vitesse du cycle infectieux	78
2.2.5	Effet de la résistance du bananier sur la dynamique parasitaire	78
2.3	Discussion	79
3	Analyse de sensibilité du modèle SiBaToka	83
3.1	Introduction	85

3.2	Matériel et méthodes	86
3.2.1	Le modèle SiBaToka	86
3.2.2	Méthodes d'analyse de sensibilité	92
3.3	Résultats	96
3.3.1	Évaluation de l'effet des composantes de résistance de la plante et de la dispersion de l'inoculum secondaire	96
3.3.2	Évaluation de l'effet des caractères agronomiques et génétiques de la plante	102
3.3.3	Évaluation du modèle complet	105
3.4	Discussion	108
3.4.1	Discussion des principaux résultats	108
3.4.2	Discussion des méthodes et du modèle	110
4	Modélisation de la dynamique d'invasion de la cercosporiose noire à l'échelle d'un territoire	115
4.1	Contextes	117
4.1.1	Contexte biologique	117
4.1.2	Contexte méthodologique	119
4.1.3	Objectifs du chapitre	120
4.2	Données	121
4.2.1	Données de présence-absence à l'échelle d'un territoire	121
4.2.2	Données environnementales	123
4.2.3	Reconstitution du paysage	125
4.3	Modèle spatio-temporel	127
4.3.1	Description du modèle développé	127
4.3.2	Simulations du modèle	129
4.4	Inférence des paramètres du modèle	133
4.4.1	Vraisemblance, loi a priori et loi a posteriori	133
4.4.2	Algorithmes et validation	134
4.4.3	Validation	137
4.5	Analyse de sensibilité du modèle	142
4.5.1	Méthode de Morris	143
4.5.2	Méthode extended FAST	144
4.6	Inférence sur données réelles : Article en préparation	148
4.7	Discussion du chapitre	162
4.7.1	Discussion des principaux résultats	162
4.7.2	Discussion méthodologique	163
	Conclusion	169
	Bibliographie	177
A	Analyse de sensibilité du modèle SiBaToka	I

Chapitre 1

Introduction

Sommaire

1.1 Épidémies végétales et invasions biologiques	3
1.1.1 Définitions	3
1.1.2 Facteurs influant sur les dynamiques épidémiques	4
Dispersion	4
Résistances	4
1.1.3 Modélisation spatio-temporelle et importance de l'échelle spatiale	5
1.2 Méthodes statistiques utilisées	5
1.2.1 Modélisation à l'échelle d'une plante	6
Les modèles mécanistes	6
Modélisation de l'hôte	6
Modélisation du pathogène	6
Dispersion à courte distance	7
1.2.2 Modélisation à l'échelle d'un paysage	7
Processus ponctuels spatio-temporels	7
Dispersion à longue distance	8
Description du paysage	8
1.2.3 Méthodes d'estimation des paramètres	8
Les méthodes Bayésiennes	9
Algorithmes d'exploration numérique MCMC	9
Test de contrôle de convergence des chaînes	11
1.2.4 Analyse de sensibilité de modèles	11
Types de méthodes	11
1.3 Notre cas d'étude : la cercosporiose noire du bananier	13
1.3.1 Le bananier	13
1.3.2 La cercosporiose noire	14
1.3.3 Le pathogène <i>Mycosphaerella fijiensis</i>	15
Biologie	15

	Cycle épidémique	16
1.3.4	Épidémiologie	16
	Conditions climatiques	17
	La croissance de l'hôte	17
	Dispersion	17
	Résistance	17
1.3.5	Origine et distribution	18
	Maladie émergente et succès d'invasion	18
1.3.6	Méthodes de lutte	18
	Pratiques culturales	18
	Lutte fongicide	19
	Variétés résistantes : amélioration et sélection variétale	19
1.3.7	Modélisation de la cercosporiose noire du bananier	19
1.4	Objectifs de la thèse	20

1.1 Épidémies végétales et invasions biologiques

1.1.1 Définitions

Une épidémie se définit comme le développement et la propagation rapide d'une maladie contagieuse, le plus souvent d'origine infectieuse, dans une population hôte.

Le développement d'une épidémie dans un agrosystème résulte des interactions entre les populations hôtes végétales, les populations pathogènes (responsables des maladies végétales) et les facteurs environnementaux (incluant les pratiques culturales). Ainsi, l'épidémiologie des maladies végétales s'attache à étudier les dynamiques de populations pathogènes au sein des populations hôtes ayant pour conséquence l'apparition d'une maladie sous l'influence de facteurs de l'environnement et de l'action de l'homme. L'épidémiologie végétale peut être considérée comme la branche écologique de la pathologie végétale.

La consommation alimentaire mondiale est devenue dépendante de quelques denrées de base que sont le blé, le riz, le maïs et la pomme de terre, mais aussi la banane, le café et le cacao qui représentent des sources d'échanges et de revenus majeures pour les pays producteurs [Anderson et al., 2004]. C'est pourquoi les maladies infectieuses causées par des virus, des champignons ou des bactéries, susceptibles d'affecter ces productions, font l'objet de recherches appliquées et théoriques. Selon la définition proposée par Anderson et al. [2004], on appelle "maladie émergente" toute maladie causée par un agent pathogène :

- (i) dont l'incidence a augmenté par expansion de son aire de répartition ou élargissement de sa gamme d'hôtes,
- (ii) dont la pathogénicité a changé,
- (iii) dont l'espèce vient d'évoluer ou bien
- (iv) que l'on vient de découvrir ou de reconnaître.

Le caractère nouveau du phénomène est commun à tous les cas d'émergence de maladies, tout comme les importants dégâts qu'elles peuvent causer. Parmi les maladies émergentes, les invasions biologiques correspondent à des introductions de pathogènes dans de nouvelles régions géographiques (on parlera dans ce cas de maladies exotiques).

Les champignons et organismes assimilés (oomycètes) sont, après les virus, les plus grands responsables de maladies émergentes de plantes. Toutes les cultures ou presque sont affectées par une ou plusieurs espèces de champignons phytopathogènes, des céréales aux arbres fruitiers (parmi lesquels les caféiers et cacaoyers), jusqu'aux plantes et arbres ornementaux. Toutes les parties des plantes peuvent être affectées, y compris fruits et graines, par une grande variété de symptômes (lésions, nécroses, chancres...). Leur impact est donc colossal en agriculture sur tous les continents, où ils peuvent réduire sévèrement les rendements et la qualité des produits agricoles, voire condamner toute production.

1.1.2 Facteurs influant sur les dynamiques épidémiques

De nombreux facteurs liés à l'agent pathogène, à l'hôte ou à l'environnement peuvent influencer sur les dynamiques épidémiques. Dans le cadre de cette thèse, nous nous intéressons aux maladies de cultures pérennes en conditions tropicales.

Une culture pérenne est une culture dont la durée de vie peut s'étaler sur plusieurs années. Ainsi, l'hôte étant présent sur une très longue période, les maladies de cultures pérennes présentent des dynamiques moins variables que celles de cultures annuelles. Dans les conditions tropicales, les conditions climatiques sont très favorables au développement des maladies. Ainsi, les dynamiques épidémiques de cultures pérennes tropicales (comme le caféier, le cacaoyer ou le bananier) ont la particularité de ne pas être chaotiques et de ne pas connaître d'extinction complète des populations pathogènes comme c'est le cas pour les maladies de milieux tempérés.

C'est pourquoi, dans ce contexte, parmi tous les facteurs influant sur les dynamiques épidémiques, nous nous focalisons sur des facteurs liés à l'agent pathogène : la dispersion ; et sur un facteur lié à l'hôte : la résistance.

Dispersion

La dispersion se définit comme le déplacement d'une propagule (spore par exemple) d'un point à un autre dans l'espace. Elle se décompose en trois étapes : l'émission de spores, le transport et le dépôt. Elle fait partie intégrante de la définition d'une épidémie causée par un pathogène et une meilleure compréhension du phénomène permet une meilleure gestion de l'épidémie [Aylor, 1999].

Résistances

Pour limiter ou empêcher le développement de pathogènes, les méthodes de contrôle des maladies consistent à perturber les dynamiques épidémiques. Il existe trois types de méthodes de lutte contre les maladies végétales :

- (i) la lutte culturale qui correspond à la mise en place de pratiques agronomiques favorables à la culture et défavorables au développement des maladies (densité de plantation, date de plantation, choix des sites de culture...)
- (ii) la lutte phytosanitaire qui consiste à perturber le cycle épidémique sur la plante (pour empêcher l'infection et/ou la reproduction du pathogène)
- (iii) la lutte génétique qui correspond à l'utilisation de variétés résistantes.

On distingue ici deux types de résistance génétique : la résistance totale qui empêche complètement le développement des pathogènes sur l'hôte et la résistance partielle, modérant uniquement leur développement. La résistance partielle réduit la sévérité de la maladie plutôt que de la supprimer [Jones, 2000]. En effet, elle permet aux hôtes de limiter ou ralentir les infections des pathogènes.

Il est utile de préciser que le terme de composante de résistance est attribué à tous les paramètres qui donnent la capacité à un hôte d'empêcher ou de réduire l'installation de pathogènes.

Une bonne compréhension des dynamiques épidémiques nécessite de prendre en compte des facteurs multiples tels ceux liés à l'agent pathogène, l'hôte et l'environnement. Les interactions étant complexes, il est difficile d'étudier les épidémies. C'est pourquoi la modélisation est un outil performant pour agréger les connaissances sur les systèmes biologiques, en particulier sur les épidémies végétales et permettre de mieux les comprendre.

1.1.3 Modélisation spatio-temporelle et importance de l'échelle spatiale

La modélisation occupe une place de plus en plus importante dans les études épidémiologiques. L'outil de modélisation permet notamment de synthétiser les processus en jeu et de simuler différents scénarios de contrôle d'une maladie. Les modèles mathématiques permettent d'approfondir les attentes des biologistes sans nécessairement avoir recours à des expérimentations.

Les questions d'échelles sont aussi cruciales dans l'étude des dynamiques épidémiques [Matthews and Haydon, 2007]. En effet, l'infection d'un hôte par un pathogène, i.e. le dépôt de l'inoculum, se concentre à une échelle spatiale évaluée au millimètre carré (Lannou et al. [2008], Mundt [2009]), et les spores produites par une reproduction rapide du pathogène peuvent être dispersées à plusieurs milliers de kilomètres (Brown and Hovmøller [2002], Aylor [2003]).

Évaluer le processus d'établissement et de propagation d'une épidémie doit donc se faire à plusieurs échelles spatiales. A l'échelle de la plante par exemple il s'agit de mieux comprendre l'infection, la croissance et la dispersion des spores sur les feuilles. A plus grande échelle spatiale, il s'agit d'évaluer globalement le développement spatio-temporel de la maladie.

Deux moyens sont usuellement employés pour modéliser la dispersion : le taux de migration et la fonction de dispersion [Ronce, 2007]. Le taux de migration définit la proportion d'individus qui se dispersent d'une population à une autre (modèle spatialement implicite).

La fonction de dispersion est une loi statistique qui décrit la probabilité qu'une spore soit déposée en un point quelconque de l'espace [Austerlitz et al., 2004] (modèle spatialement explicite). Cette densité de probabilité peut inclure tout un ensemble de relations qu'entretient la dispersion avec l'espace, dont la plus communément admise est la décroissance de la probabilité d'être déposé avec la distance à la source.

La résistance se modélise en général en incluant des variables d'entrée des modèles définissant les niveaux de résistances [Lô Pelzer et al., 2010b].

1.2 Méthodes statistiques utilisées

Nous décrivons dans un premier temps les différentes techniques de modélisation utilisées selon les échelles spatiales d'étude puis nous détaillons les méthodes d'inférence utilisées dans la thèse et pour finir nous présentons les différentes méthodes d'analyses de sensibilité disponibles et plus précisément celles utilisées par la suite.

1.2.1 Modélisation à l'échelle d'une plante

A l'échelle d'une plante, de nombreux auteurs se sont intéressés aux pathogènes responsables des dégâts observés dans les principales cultures, dans le but de mieux comprendre leurs dynamiques spécifiques ou en interactions avec l'hôte, afin d'améliorer les méthodes de lutte [Jeger, 2004].

En effet, parmi ces modèles certains décrivent spécifiquement la dynamique du pathogène ou de la plante et d'autres couplent les deux dynamiques. Il existe de nombreuses manières de décrire la dynamique d'un pathogène sur une plante, la technique utilisée étant basée sur la question de recherche et la spécificité du pathogène. Par exemple, Papaïx et al. [2011] utilisent des méthodes "State-Space-Model" pour décrire l'interaction entre un pathogène et des hôtes de sensibilités différentes. Ou encore, [Véra, 2004] s'intéresse à la modélisation de la croissance des plantes par des modèles linéaires mixtes.

Les modèles mécanistes

Nous nous intéressons ici plus particulièrement aux modèles les plus utilisés dans la littérature, les modèles mécanistes. Ces modèles cherchent à décrire précisément les mécanismes mis en jeu dans la dynamique de l'hôte ou du pathogène en intégrant une représentation discrète de l'espace et du temps [De Wolf and Isard, 2007]. Ils permettent donc de prendre en compte de manière détaillée les différentes étapes de la croissance de l'hôte et du développement du pathogène.

Modélisation de l'hôte

Les modèles décrivant la dynamique d'une plante en croissance s'intéressent principalement à décrire les stades de développement de la plante, en particulier le développement foliaire et la croissance des feuilles principalement dans le but d'évaluation de méthode de lutte. Par exemple Milne et al. [2003] proposent un modèle décrivant la croissance d'une vigne selon ses stades de développement dans le but de fournir un outil d'aide à la décision dans l'application de fongicide sur les feuilles. [Tixier et al., 2004] développent un modèle de croissance du bananier basé sur deux composantes, le développement piloté par le temps physiologique et l'accumulation de biomasse dans le but de reproduire fidèlement le système de culture et de pouvoir y intégrer des dynamiques de ravageurs [Tixier et al., 2006].

Modélisation du pathogène

Les modèles mécanistes décrivent précisément les différentes étapes du cycle infectieux des différents pathogènes, soit de manière générique (Caubel et al. [2012], Casadebaig et al. [2012]) ; modèles s'adaptant à tout type de pathogène, soit en tenant compte des particularités de chaque pathogène étudié (Calonnec et al. [2008], Audsley et al. [2005], Rossi et al. [2008]). Afin de tenir compte de l'existence d'une forte dépendance entre la dynamique de l'hôte et la transmission du pathogène [Costes et al., 2013] les étapes du cycle infectieux des pathogènes sont décrites dans ces modèles sur un hôte en croissance.

Que le modèle soit générique ou non, les principales étapes de la dynamique épidémique sont au nombre de quatre : l'infection (arrivée des spores sur la plante), la croissance des lésions, la reproduction (production de spores) et la dispersion.

Dispersion à courte distance

A l'échelle d'une plante, la dispersion peut être modélisée soit de manière mécaniste ou bien en utilisant des fonctions de dispersion. Par exemple, Willocquet and Savary [2004] définissent trois dispersions distinctes : une auto-déposition de spores, une dispersion à l'échelle de la feuille et une dispersion à l'échelle de la plante suivant des taux de redépôt définis. Ou au contraire, Caillonnet et al. [2008] modélisent la dispersion de l'oïdium de la vigne par une fonction de dispersion particulière définie à l'intérieur d'un cône d'angle fixé et dont l'axe est orienté dans la direction du vent. La fonction décroît exponentiellement avec la distance axiale à la source et linéairement en fonction de position radiale dans le cône.

1.2.2 Modélisation à l'échelle d'un paysage

La construction de modèles spatio-temporels en épidémiologie se base en partie sur le type de données disponibles pour l'étude. On peut distinguer deux classes de modèles, ceux s'intéressant à l'occupation d'un territoire par une espèce utilisant des données quantitatives (par exemple des quantités de maladies [Soubeyrand et al., 2008]) et ceux modélisant l'occupation d'un territoire par une espèce en utilisant des données d'occurrence (par exemple la présence-absence d'un individu [Facon and David, 2006]).

Processus ponctuels spatio-temporels

A l'échelle d'un paysage, les objectifs sont de comprendre selon quel processus et quels déterminants une maladie apparaît en différents sites et à différentes dates. Pour cela, sachant que la propagation d'une épidémie peut être considérée comme un processus stochastique [Aylor, 2003], l'utilisation de processus ponctuels spatio-temporels semble bien adaptée comme le montrent Vaillant et al. [2010] et Brown and Hovmøller [2002] dans l'analyse de la dynamique de ravageurs de la canne à sucre, ou encore Soubeyrand et al. [2009] qui modélisent la propagation du pathogène *Plantago lanceolata* au moyen d'un processus ponctuel de Poisson spatio-temporel.

Un processus ponctuel est défini comme une mesure stochastique de dénombrement modélisant les localisations de points dans un espace métrique. La loi d'un processus ponctuel, sous certaines conditions de régularité, est complètement caractérisée par son processus d'intensité conditionnelle noté $\lambda(\cdot)$ [Daley and Vere-Jones, 2003].

En général, l'intensité $\lambda(\cdot)$ varie au cours du temps et/ou dans l'espace et dépend des événements antérieurs, et l'on modélise un processus ponctuel spatio-temporel par son intensité stochastique conditionnelle. Celle-ci désigne le taux espéré d'occurrences d'événements autour d'un site x présent dans la région spatiale d'étude X , conditionnellement à l'histoire du processus jus-

qu'à la date t exclue ($t \in [0, T]$ où $[0, T]$ est la période d'étude), on note l'intensité de ce processus :

$$\lambda(x, t|H_{t-})$$

où H_{t-} représente l'historique des observations jusqu'à la date t exclue.

Pour des données de la forme $(x_i, t_i)_{i \in [1, n]}$, couples de positions et dates d'occurrence, correspondant à la réalisation d'un processus ponctuel spatio-temporel simple (processus dans lequel au plus une occurrence se produit à un instant donné) dans l'espace X au cours de la période $[0, T]$, la log-vraisemblance se note (Daley and Vere-Jones [2003][Diggle, 2006]) :

$$L(\theta; (x_i, t_i)) = \sum_{0 < t_j \leq T} \log(\lambda_\theta(x_j, t_j|H_{t_j-})) - \int_0^T \int_X \lambda_\theta(x, t|H_{t-}) dx dt \quad (1.1)$$

où $\lambda_\theta(x_j, t_j|H_{t_j-})$ est l'intensité du processus en (x_j, t_j) et H_{t-} représente l'historique des observations jusqu'à la date t exclue, θ est le paramètre (ou vecteur de paramètres) sur lequel porte l'inférence statistique.

Dispersion à longue distance

La dispersion à longue distance est un phénomène observé chez la plupart des espèces végétales (Austerlitz et al. [2004], Schurr et al. [2008]). Les fonctions de dispersion peuvent être classées en trois catégories, basées sur la vitesse de décroissance de la fonction pour des distances infinies : i) les fonctions à décroissance exponentielle, ii) les fonctions à décroissance plus rapide que n'importe quelle exponentielle (dits à queues légères) et iii) les fonctions à décroissance plus lente que n'importe quelle exponentielle (dits à queues lourdes).

Description du paysage

Le paysage peut être décrit en considérant l'espace comme homogène ([Fewster, 2003], [Long et al., 2012] [Ross, 2011]) ou hétérogène (Schreiber and Lloyd-Smith [2009], Meentemeyer et al. [2011], Cook et al. [2007]).

Sachant l'influence considérable de l'hétérogénéité du paysage sur une invasion [Real and Biek, 2007], et l'effet des variables du paysage sur la dynamique de propagation d'une maladie [Ostfeld et al., 2005] (notamment sur les dynamiques de maladie des plantes [Plantegenest et al., 2007]), nous allons considérer dans le modèle spatio-temporel développé au Chapitre 4 l'invasion d'un pathogène en paysage hétérogène.

1.2.3 Méthodes d'estimation des paramètres

Il existe un large panel de méthodes pour estimer les paramètres d'un modèle complexe. Ces méthodes peuvent être regroupées en deux principales approches : l'approche fréquentiste et l'approche bayésienne [Makowski et al., 2006a].

Les principales différences entre ces deux méthodes concernent d'une part les données et d'autre part les paramètres. En effet, l'approche fréquentiste n'utilise qu'un jeu de données d'ob-

servations alors que la méthode Bayésienne inclut en plus des informations *a priori* sur les paramètres. Dans les méthodes bayésiennes, les paramètres du modèle sont considérés comme des variables aléatoires alors que dans l'approche fréquentiste les paramètres ont des valeurs fixes, à estimer.

Les avantages des méthodes bayésiennes sont donc de prendre en compte des informations *a priori* sur les paramètres et d'obtenir des distributions de probabilité des valeurs des paramètres *a posteriori*. Ces méthodes constituent un outil puissant pour l'estimation de la variabilité dans les processus écologiques [Clark, 2005] et sont donc de plus en plus utilisées lorsque les paramètres ont un sens biologique ou physique [Makowski et al., 2006a].

Les méthodes Bayésiennes

Les statistiques bayésiennes s'appuient sur la généralisation de la formule de Bayes :

$$\pi(\theta|Y) = \frac{\pi(Y|\theta)\pi(\theta)}{\int_{\theta} \pi(Y|\theta)\pi(\theta)d\theta} \quad (1.2)$$

où Y est le vecteur d'observations que l'on doit expliquer à partir d'un modèle faisant intervenir un paramètre θ , $\pi(\theta|Y)$ est la distribution *a posteriori* du paramètre θ , $\pi(\theta)$ est la distribution *a priori* de θ intégrant la connaissance *a priori* du paramètre, $\pi(Y|\theta)$ est la vraisemblance du modèle et $\int_{\theta} \pi(Y|\theta)\pi(\theta)d\theta$ est la constante de normalisation.

La constante de normalisation est souvent impossible à calculer de façon analytique, on a alors :

$$\pi(\theta|Y) \propto \pi(Y|\theta)\pi(\theta) \quad (1.3)$$

Les paramètres de la distribution *a posteriori* sont fonctions des paramètres de la distribution *a priori*, de la vraisemblance et des données. Cette densité *a posteriori* peut être utilisée à diverses fins comme estimer la valeur la plus probable d'un paramètre (en considérant le mode ou la moyenne), calculer l'incertitude sur l'estimation des paramètres ou encore calculer l'incertitude sur les variables simulées par le modèle à partir de l'incertitude sur l'estimation des paramètres.

Lorsque les lois de la vraisemblance et la distribution *a priori* ne sont pas conjuguées (Parent and Bernier [2007] chapitres 7 et 8), des algorithmes permettent d'échantillonner des valeurs de θ dans la distribution *a posteriori*.

Algorithmes d'exploration numérique MCMC

Parmi les algorithmes capables d'explorer la distribution *a posteriori* des paramètres, on peut citer les algorithmes de Monte Carlo par Chaîne de Markov (MCMC), les algorithmes d'échantillonnage pondéré (importance sampling en anglais) ou bien certaines méthodes de filtrage (voir [Robert and Casella, 2010] pour le détail de toutes ces méthodes).

Dans la suite, nous nous intéresserons principalement aux méthodes MCMC car ces méthodes, les plus utilisées, permettent le calcul bayésien d'une très large variété de modèles [Parent and Bernier, 2007]. Ces méthodes permettent de d'obtenir un échantillon de la loi *a posteriori* d'intérêt, en générant une chaîne de Markov qui tend vers cette loi. Autrement dit, ces méthodes s'appuient

sur le comportement limite stationnaire de la chaîne, qui parcourt l'espace des états possibles de θ avec une fréquence convergeant vers la loi *a posteriori* $\pi(\theta|Y)$ (basé sur le théorème d'ergodicité de la chaîne de Markov générée). Il est à noter que la vitesse de passage à ce comportement stationnaire peut être très lente.

Les méthodes MCMC génèrent une suite de variables aléatoires $(\theta^1, \dots, \theta^n, \dots)$ et hormis la première à laquelle on donne une valeur arbitraire, chacune d'entre elles dépend uniquement de celle qui précède.

Nous nous intéresserons ici plus particulièrement aux algorithmes de Métropolis-Hastings et de Gibbs. L'algorithme général de Métropolis-Hastings peut se synthétiser de la manière suivante :

Algorithme 1 Algorithme général de Metropolis-Hastings

Initialisation du paramètre : $\theta^{[0]} \sim \pi(\theta)$

Itération $m = 1$.

tant que $m \leq M$ **faire**

$\theta^* \sim q(\cdot, \theta^{[m-1]})$

$$\rho = \frac{\pi(\theta^*|Y) q(\theta^{[m-1]}, \theta^*)}{\pi(\theta^{[m-1]}|Y) q(\theta^*, \theta^{[m-1]})}$$

Soit u , tel que $u \sim \mathcal{U}[0, 1]$

si $u \leq \rho$ **alors**

$\theta^{[m]} = \theta^*$

sinon $\theta^{[m]} = \theta^{[m-1]}$

fin du si

$m = m + 1$

fin du tant que

Sortie(s) $(\theta^{[1]}, \dots, \theta^{[M]})$

Après initialisation des paramètres avec les lois *a priori*, ils sont mis à jour à chaque itération en se basant sur une loi de probabilité appelée *loi de proposition*, notée $q(\cdot, \cdot)$ dans l'algorithme. La force des sauts aléatoires est réglée par la variance de cette loi de proposition. Cette dernière est généralement caractérisée par une loi normale (ou normale tronquée) centrée sur la valeur du paramètre de l'itération précédente et dont la variance est fixée par l'utilisateur (O'Hara et al. [2002], Cook et al. [2007]).

La performance de l'algorithme de Metropolis-Hastings peut être évaluée en calculant le taux d'acceptation des paramètres proposés par Gelman et al. [2004].

L'échantillonnage de Gibbs ne demande pas de mettre en place une fonction d'exploration de l'espace d'états des paramètres. Cette méthode exige la possibilité de simulation des conditionnelles complètes de la loi d'une composante de θ sachant toutes les valeurs fixées des $k - 1$ autres composantes.

En considérant un vecteur θ à k composantes, l'algorithme général de Gibbs peut se synthétiser de la manière suivante :

Algorithme 2 Algorithme général de Gibbs

Initialisation du paramètre : $\theta_0^{[1]} \sim \pi(\theta)$

Itération $m = 1$.

tant que $m \leq M$ **faire**

Générer $\theta_{m+1}^{[1]}$ en simulant selon la loi $\pi(\theta^{[1]} | \theta_m^{[2]}, \dots, \theta_m^{[j]}, \dots, \theta_m^{[k]})$

Générer $\theta_{m+1}^{[2]}$ en simulant selon la loi $\pi(\theta^{[2]} | \theta_{m+1}^{[1]}, \theta_m^{[3]}, \dots, \theta_m^{[k]})$

Générer $\theta_{m+1}^{[j]}$ en simulant selon la loi $\pi(\theta^{[j]} | \theta_{m+1}^{[1]}, \dots, \theta_{m+1}^{[j-1]}, \theta_m^{[j+1]}, \dots, \theta_m^{[k]})$

Générer $\theta_{m+1}^{[k]}$ en simulant selon la loi $\pi(\theta^{[k]} | \theta_{m+1}^{[1]}, \dots, \theta_{m+1}^{[k-1]})$

$m = m + 1$

fin du tant que

Sortie(s) $\theta^M = \{\theta^{[k]}\}$

L'étape m est scindée en k sous-étapes. Chaque sous-étape est un tirage dans la distribution conditionnelle complète de chaque $\theta^{[j]}$ connaissant l'ensemble des autres variables. Un nouveau vecteur θ_{m+1} est ainsi généré à partir des éléments de θ_m .

Test de contrôle de convergence des chaînes

La méthode la plus utilisée pour contrôler que la phase ergodique est atteinte est celle de Gelman-Rubin [Gelman and Rubin, 1992]. Cette méthode est basée sur S trajectoires de longueur n initiées indépendamment (on choisit généralement $S = 3$). Ce critère a pour expression

$$\hat{R} = \frac{(n-1)/nB + 1/nA}{B}$$

où A est la variance inter-suite et B est la variance intra-suite. Ce rapport tend vers la valeur 1 si la convergence est atteinte. Ce critère est calculé par exemple dans le package CODA du logiciel R [R Development Core Team, 2008].

1.2.4 Analyse de sensibilité de modèles

Types de méthodes

L'analyse de sensibilité permet de hiérarchiser l'importance des différents paramètres d'entrée d'un modèle. Les résultats d'une analyse de sensibilité peuvent ainsi aider à identifier les paramètres qu'il serait utile de connaître plus précisément, soit par le biais d'estimations statistiques soit par des expérimentations.

Deux types d'analyses de sensibilité d'un modèle sont réalisables, l'analyse locale qui évalue la variabilité en un point donné et l'analyse globale qui s'intéresse à l'ensemble du domaine de variation des variables d'entrée. Nous nous intéresserons par la suite aux différentes méthodes d'analyses globales.

Deux grandes classes de méthodes d'analyses globales sont distinguées : les méthodes de criblage (tri grossier des entrées les plus influentes parmi un grand nombre) et les méthodes basées sur les mesures d'importance (indices quantitatifs donnant l'influence de chaque entrée sur la variance de la sortie) [Iooss, 2010].

Parmi les méthodes de criblage, les principales méthodes utilisées sont la méthode d'analyse de sensibilité de Morris et l'analyse de la variance (ANOVA) sur des plans factoriels. Ces méthodes s'appuient sur une discrétisation du domaine des variables d'entrée du modèle selon une grille régulière, obtenue en discrétisant le domaine de variation de chaque variable d'entrée. Nous présentons ici en détail la méthode de Morris utilisée dans le Chapitre 3.

L'objectif de la méthode de Morris est de déterminer quelles variables d'entrée peut être considérée comme ayant des effets : négligeables, linéaires ou additifs, ou bien non linéaires ou impliqués dans des interactions avec d'autres variables [Campolongo et al., 2007]. La méthode utilise le principe d'un plan d'expérience One-At-a-Time (OAT), qui consiste à faire varier la valeur d'une variable d'entrée en laissant les autres à une valeur fixe afin de détecter l'effet de ce seul paramètre sur la sortie du modèle, sur différentes combinaisons de valeurs des différents facteurs.

Soit $f(X)$ la sortie d'intérêt du modèle dans laquelle $X = (x_1, \dots, x_i, \dots, x_N)$ est une combinaison de valeurs des N variables d'entrée, l'effet élémentaire de la i^{eme} variable est

$$EE_i(X) = \frac{f(x_1, \dots, x_i + \Delta, \dots, x_N) - f(x_1, \dots, x_i, \dots, x_N)}{\Delta}$$

avec $\Delta \in \{1/(p-1), \dots, 1 - 1/(p-1)\}$ et p est le nombre de niveaux de discrétisation des domaines de variation des variables d'entrée.

En construisant r différents scénarios (combinaisons de valeurs pour les variables d'entrée), balayant toutes les valeurs possibles pour les différentes variables d'entrée, on obtient pour chaque variable une distribution des valeurs représentant son effet dans les différents scénarios.

Ces distributions sont caractérisées par la moyenne (μ_n^*) et la variance (σ_n^2) des $EE_i(X)$.

$$\mu_n^* = \frac{\sum_{i=1}^r |EE_i(X)|}{r}$$

$$\sigma_n^2 = \frac{\sum_{i=1}^r [EE_i(X) - 1/r \sum_{i=1}^r EE_i(X)]^2}{r-1}$$

Une variable possédant une forte moyenne μ_n^* sera considérée comme ayant une forte importance sur la sortie du modèle, une forte valeur de la variance σ_n^2 sera indicatrice d'une interaction entre la variable considérée et une autre ou alors d'un effet non linéaire de la variable.

Parmi les méthodes basées sur les mesures d'importance, il existe principalement des méthodes de régression linéaire et des méthodes basées sur l'analyse de variance utilisant une décompo-

sition de la variance fonctionnelle. Nous nous intéresserons plus particulièrement ici à ces dernières.

Ces méthodes permettent d'estimer l'importance des variables d'entrée sur la sortie du modèle $Y = f(X)$, où $X = (X_1, \dots, X_N)$ est le vecteur des variables d'entrée du modèle, en utilisant la décomposition de la sortie en somme de fonctions élémentaires :

$$V(Y) = \sum_{i=1}^N V_i + \sum_{1 \leq i < j \leq N} V_{ij} + \dots + V_{1,2,\dots,N}$$

où $V_i = V[E(Y|X_i)]$ est l'effet principal du paramètre X_i sur la sortie d'intérêt Y et les autres termes représentent les interactions.

Les indices de Sobol, ou "mesures d'importance basées sur la variance" [Iooss, 2010] sont :

$$S_i = \frac{V_i}{V(Y)}$$

$$ST_i = \frac{V(Y) - V_{-i}}{V(Y)}$$

où V_{-i} est la somme de tous les termes n'incluant pas i , S_i représente l'indice de sensibilité principal du paramètre i et ST_i représente l'indice de sensibilité total du paramètre i (somme de tous les effets du paramètre).

Les deux indices sont compris entre 0 et 1, et sont interprétables en terme de proportion de variance expliquée par le paramètre. De faibles valeurs de ces indices indiquent des effets négligeables et inversement.

Pour estimer ces indices de sensibilité, des méthodes basées sur des échantillons de Monte Carlo ont été développées (méthode de Sobol) mais sont très coûteuses en temps de calcul, nous nous sommes concentrés sur la méthode e-FAST (extended Fourier Amplitude Sensitivity Test) moins coûteuse en temps de calcul et tout aussi performante (Varella et al. [2010], Saltelli et al. [1999], Makowski et al. [2006b]). La difficulté principale dans le calcul de ces indices est qu'ils requièrent le calcul d'intégrales en grandes dimensions. La méthode e-FAST utilise un échantillonnage déterministe pour explorer l'espace des paramètres ce qui rend possible de réduire ces intégrales à une seule dimension par le biais d'une décomposition de Fourier (les détails de cette décomposition sont disponibles dans (Faivre et al. [2013], Confalonieri et al. [2010b], Wang et al. [2013])).

1.3 Notre cas d'étude : la cercosporiose noire du bananier

1.3.1 Le bananier

Le bananier est une plante monocotylédone appartenant à la famille des Musacées (*Musa spp.*). Le bananier est une herbe géante à croissance rapide pouvant atteindre plus de 15 mètres de hauteur. Ses larges feuilles, d'environ 1 m², s'enroulent les unes sur les autres et forment un pseudo-tronc donnant à cette plante un aspect d'arbre.

Comme cette plante produit annuellement un régime constitué de fruits stériles (sans graines), la pérennité du bananier est assurée par sa reproduction végétative. En effet, avant la récolte, la plante émet des ramifications latérales (rejets) qui vont lui succéder. Chaque rejet, provenant de sa tige souterraine, peut donner un plant de bananier qui peut succéder au pied-mère sur la même souche ou être détaché pour être planté et cultivé ailleurs.

Il existe une grande diversité génétique au sein de cette famille botanique avec plus de 1000 variétés décrites. Les bananiers cultivés (dits cultivars) se divisent en deux grands sous-groupes : celui des bananes douces ou bananes dessert et celui des bananes à cuire, parmi lesquelles les plantains occupent une place prépondérante. Dans les pays tropicaux, les variétés à cuire constituent un aliment de base pour plus de 400 millions de personnes.

En effet, le bananier est classé au quatrième rang économique mondial des cultures destinées à l'alimentation. Il est cultivé dans plus de 130 pays principalement dans les régions tropicales et sub-tropicales du monde. La production mondiale est estimée à 130 millions de tonnes (FAO, 2009) dont 16 millions sont destinées à l'exportation, dont 2,7 millions sont produites dans les Caraïbes et environ 300000 tonnes aux Antilles françaises.

La banane pour l'exportation, concernant surtout les bananes dessert, est cultivée dans de vastes plantations industrielles dont les principaux producteurs sont l'Équateur, le Costa-Rica, les Philippines où les rendements moyens vont de 40 à 60 tonnes /hectares alors que la banane d'autoconsommation est cultivée en majorité dans de petites fermes familiales, appelées aux Antilles "jardin créoles", ayant de faibles rendements (1 à 12 tonnes par an).

1.3.2 La cercosporiose noire

Parmi les divers parasites auxquels doivent faire face ces cultures, la cercosporiose noire ou maladie des raies noires (Black Leaf Streak Disease en anglais) est la maladie foliaire du bananier la plus grave car la plus destructrice et la plus coûteuse en termes de méthodes de lutte [Churchill, 2011].

Causée par le champignon phytopathogène *Mycosphaerella fijiensis*, cette maladie foliaire entraîne une défoliation sévère (Figure 1.1). La capacité photosynthétique des feuilles est amoindrie et provoque une baisse de qualité et de quantité de fruits récoltés, entraînant des pertes de rendement allant de 20% à 100% (Castelan et al. [2012], Marin et al. [2003], Jones [2000]). En effet, la maladie entraîne non seulement une diminution du poids des fruits mais également accélère la maturité des fruits (réduit la durée de vie verte des fruits). Cette maturité précoce des fruits est le dégât le plus problématique pour les productions industrielles destinées à l'exportation pour lesquelles les durées de transport entre les zones tropicales de production et les pays importateurs européens, japonais ou américains, exigent des durées de vie vertes très longues (plus de 3 semaines) [Marin et al., 2003].



FIGURE 1.1 – Défoliation du bananier causée par la cercosporiose noire : a) à l'échelle de la plante, b) à l'échelle d'une feuille. (Photos : Catherine Abadie)

1.3.3 Le pathogène *Mycosphaerella fijiensis*

Biologie

Mycosphaerella fijiensis est un champignon appartenant à la classe des ascomycètes (champignon dont les spores se forment dans des asques) hétérothalliques (dotés de thalles de type sexuels opposés). Il possède deux types de reproduction, et émet donc deux types de spores : les conidies, issues de la reproduction asexuée, et les ascospores, issues de la reproduction sexuée par rencontre de deux lésions de types sexuels opposés (Figure 1.2). Les populations pathogènes responsables de la cercosporiose noire présentent des niveaux de diversité génétique et un potentiel adaptatif élevés (Carlier et al. [1996], Churchill [2011]).



FIGURE 1.2 – Spores produites par le pathogène *Mycosphaerella fijiensis* : a) conidies (spores asexuées), b) ascospores (spores sexués). (Photos : Cirad)

Cycle épidémique

Le cycle épidémique *Mycosphaerella fijiensis* peut être décrit de la manière suivante (Figure 1.3) : une fois les spores déposées sur les feuilles, elles germent et pénètrent par les stomates. Après une durée d'incubation allant de 10 à 30 jours, les premières lésions vont apparaître puis croître. Après une période de latence conidienne, la reproduction asexuée se met en place et des conidies (spores asexuées) sont produites à la surface de la feuille. Ces spores proviennent de conidiophores à la surface des feuilles, et vont être dispersées essentiellement par le ruissellement de l'eau sur les feuilles. Après la sporulation conidienne, les lésions continuent de croître et la rencontre de deux lésions de types sexuels opposés donne lieu à la reproduction sexuée dans des périthèces dans l'épaisseur des feuilles. Cette reproduction donne lieu à des plages nécrotiques provoquées par la coalescence des lésions. Après pénétration d'eau (pluie, rosée) dans les périthèces, les ascospores (spores sexuées) sont expulsées de façon active dans l'air. Ces spores expulsées vont se disperser principalement dans l'air Churchill [2011]).

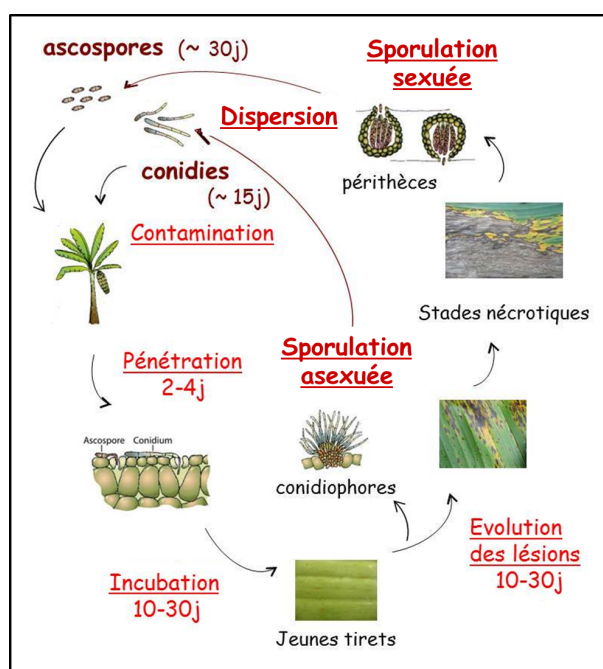


FIGURE 1.3 – Cycle épidémique de la cercosporiose noire du bananier [Rieux, 2011].

1.3.4 Épidémiologie

La cercosporiose noire est une maladie polycyclique (une dizaine de cycles par an). Deux types d'infections sont définies : l'infection primaire ou allo-infection correspond au dépôt de spores venant de l'extérieur sur des feuilles saines et l'infection secondaire résulte de l'auto-contamination de la plante.

Conditions climatiques

L'humidité relative (nécessaire à la pénétration du pathogène et à sa sporulation) est le principal facteur climatique influant sur le cycle infectieux nécessaire à l'infection et à la sporulation. D'autres facteurs climatiques peuvent jouer sur le cycle infectieux, en particulier la température inférieure à 20 °C et l'intensité lumineuse (ombrage) en le ralentissant [Marin et al., 2003]. Les zones tropicales sont donc des zones privilégiées pour le développement du pathogène. Dans ces conditions optimales, les niveaux de populations pathogènes sont très élevés lorsque les hôtes sont sensibles à la maladie.

La croissance de l'hôte

Les bananiers sont caractérisés par une émission foliaire quasi hebdomadaire qui se traduit donc par un décalage chronologique entre chaque rang de feuille. Bien que les infections primaires d'un plant puissent potentiellement avoir lieu sur toutes les feuilles, la très grande majorité d'entre elles surviennent pendant le déroulement de chaque feuille (stade cigare) au sommet du bananier [Gauhl, 1994]. Par conséquent, il existe un gradient d'évolution de la maladie entre les jeunes feuilles du haut et les vieilles feuilles situées en bas de la plante. Enfin contrairement à d'autres plantes, toutes les feuilles du bananier (de la première à la dernière avant floraison) sont réceptives à la maladie.

Dispersion

La dispersion des conidies se produit principalement par la pluie et est considérée comme responsable de l'auto-contamination de la plante [Marin et al., 2003]. Rieux et al. [2014] estiment leur distance moyenne de dispersion à 3 mètres de la source infectieuse.

La distance de dispersion des ascospores, spores responsables de la contamination à longue distance et dispersées uniquement par le vent, a été étudiée à différentes échelles spatiales, suivant des techniques d'expérimentations au champ (Rutter et al. [1998], Burt et al. [1998]) et en utilisant des techniques de marquage génétique afin d'évaluer de manière plus précise cette dispersion. Rieux et al. [2014] estiment que la distance moyenne de dispersion des ascospores (toutes directions confondues) est évaluée à 283 mètres et les distances maximales relevées correspondent à des lésions observées aux bornes de l'expérimentation (1 km de la source). Amil et al. [2007] détectent la présence du pathogène à plus de 6 km de la zone de production.

Résistance

La majorité des cultivars en particulier les plantains et les bananes dessert pour l'exportation (groupe Cavendish) sont sensibles à la cercosporiose noire. Cependant, il existe des variétés et cultivars naturels résistants à la maladie. Fouré et al. [2000] a suggéré l'existence de deux types de résistance : une résistance totale (qui empêcherait toute reproduction du pathogène) et une résistance partielle qui ralentirait le cycle épidémique. Les mécanismes inhérents aux deux types de résistance ne sont pas encore élucidés.

Ces cultivars n'ont jamais été utilisés pour la production d'exportation car cela obligerait l'ensemble de la filière à modifier de nombreux procédés post-récolte actuellement en place et spécifiquement adaptés à la culture des bananiers de type Cavendish, banane "dessert" [De Lapeyre de Bellaire et al., 2010].

1.3.5 Origine et distribution

Maladie émergente et succès d'invasion

La cercosporiose noire a été décrite pour la première fois à Fidji en 1963. Il a été montré que son aire d'origine serait l'Asie du sud-Est (Carlier et al. [1996], Robert et al. [2012]). En deux décennies, elle a envahi les zones bananières tropicales et subtropicales de tous les continents : elle a été détectée sur le continent africain en 1978 au Gabon et sur le continent américain en 1972 au Honduras [Jones, 2000]. Par une approche de phylogéographie, Robert et al. [2012] suggèrent des histoires d'invasion différentes en Afrique (avec une introduction unique) et en Amérique latine avec plusieurs introductions indépendantes.

Dans l'archipel des Caraïbes, elle a été détectée pour la première fois à Cuba fin 1989, la source d'introduction probable étant le Honduras. Depuis les années 1990, elle progresse suivant deux fronts de propagation : un Ouest-Est de Cuba à Porto Rico où elle a été détectée en 2004 et un Sud-Nord depuis Trinidad où elle a été détectée en 2003 [Fortune et al., 2005] jusqu'aux Petites Antilles. Aux Antilles françaises, la première détection date de 2010 en Martinique [Ioos et al., 2011] et de 2012 en Guadeloupe [DAAF, communiqué du 24/01/2012].

La maladie est ainsi présente dans plus de 100 pays, à part dans quelques pays subtropicaux comme l'Australie, les Canaries, l'Afrique du sud et certains pays d'Afrique du nord. C'est un exemple d'émergence mondiale contemporaine réussie.

1.3.6 Méthodes de lutte

Comme énoncé précédemment, la maladie ne s'attaque pas directement aux fruits mais les dommages sont considérables à la récolte. Il s'avère donc indispensable de lutter contre la cercosporiose noire lorsque la production est vouée à l'exportation, les variétés cultivées du groupe des Cavendish étant sensibles à la maladie.

Pratiques culturales

Afin de réduire la quantité d'inoculum présent sur la plante ou dans une parcelle, il est recommandé de pratiquer une élimination régulière des feuilles nécrosées [Marin et al., 2003]. Il est aussi recommandé de maintenir une bonne croissance des bananiers par une alimentation en eau optimale, une fertilisation régulière et une élimination des adventices qui maintiennent une humidité élevée dans la parcelle favorable à la maladie [Lassoudière, 2006].

Lutte fongicide

La principale lutte actuelle dans les productions destinées à l'export est la lutte chimique. Il s'agit d'applications régulières par voie aérienne de fongicides systémiques mélangés à de l'huile minérale et, bien que soumise à de nombreuses réglementations, son impact sur l'environnement a remis en cause ces techniques dans les Antilles Françaises : depuis octobre 2013, l'application de fongicide par voie aérienne est interdite.

Variétés résistantes : amélioration et sélection variétale

Une des alternatives à la lutte chimique est l'utilisation de variétés résistantes mais il n'existe pas encore de variétés commerciales résistantes à la cercosporiose noire.

Différents programmes de création variétale (dont celui du CIRAD) ont été mis en place à partir des années 1990. Des hybrides de bananes dessert pour l'export ont été créés et sélectionnés au CIRAD en Guadeloupe dans le cadre du Plan Banane Durable mis en place en 2009 et soutenu par le ministère de l'agriculture. Les premiers hybrides sont actuellement en expérimentation et en phase de sélection en station et chez les producteurs. Mais aucune variété résistante ou partiellement résistante n'est cultivée à grande échelle aux Antilles. La sélection de ces variétés hybrides, en d'autres termes, l'évaluation des résistances partielles de ces hybrides pose des contraintes dans le temps et dans l'espace ; un cycle de culture étant d'environ un an et un bananier occupant environ 6 m².

On se demande alors, comment aider à la sélection et au déploiement des variétés hybrides créées ?

1.3.7 Modélisation de la cercosporiose noire du bananier

La modélisation nous apparaît comme un outil adéquat pour répondre à la question précédente. La littérature ne fournit pas de modèle d'interaction bananier/cercosporiose noire, hormis une courte communication à congrès [Arneson, 1989] où le modèle présenté est créé pour simuler la cercosporiose jaune sur un bananier. Il s'agit d'un modèle mécaniste où la production de feuilles, l'infection, le développement des lésions, la sporulation du champignon (asexuée et sexuée), la dispersion des spores et la réaction aux fongicides sont modélisés. Aucun document plus complet de ce modèle n'est cependant disponible.

Concernant la modélisation du bananier, Tixier et al. [2004] dans le modèle SIMBA-POP décrivent précisément le système racinaire du bananier et modélisent la production de fruits. Dans ce modèle, chaque plante se développe à son rythme sur plusieurs cycles de culture mais les caractéristiques propres à la croissance du bananier ne sont cependant pas prises en compte (changement de rang de feuilles, formation du cigare, nombre de feuilles maximum par plante, etc.).

1.4 Objectifs de la thèse

Cette thèse a pour objectif de conceptualiser et de mettre en place des outils de modélisation de la cercosporiose noire du bananier permettant de prendre en compte divers facteurs pouvant avoir une influence sur la dynamique de la maladie à différentes échelles spatiales.

A l'échelle de la plante, [Abadie et al. \[2009\]](#) ont montré que les composantes de résistance influent sur les étapes du cycle épidémique du pathogène. On se demande quels sont les mécanismes les plus importants mis en jeu dans le développement de la maladie sur une plante ? Quelles sont alors les composantes de résistance partielle les plus efficaces ? Nous allons donc nous attacher à décrire la dynamique épidémique de la cercosporiose noire en modélisant les étapes du cycle infectieux. L'objectif ici est de mieux comprendre, d'évaluer et d'explorer la dynamique de la cercosporiose noire et d'évaluer l'efficacité des composantes de résistance.

A l'échelle d'un territoire, [Papaïx et al. \[2011\]](#) par exemple, ont montré que l'aménagement des variétés résistantes et les variables climatiques peuvent réduire la pression parasitaire. On se demande ici, quels sont les types de paysage ayant le plus d'impact sur la dynamique de la maladie ? Quelle peut être la distance de dispersion des spores à cette échelle sachant la composition du paysage ? Nous allons ici nous attacher à décrire la dynamique spatio-temporelle de la cercosporiose noire en fonction des caractéristiques paysagères et climatiques d'un territoire, plus précisément d'une île des Antilles Françaises ; la Martinique. L'objectif est d'analyser, en fonction des caractéristiques du paysage, l'invasion de la cercosporiose noire pour mieux appréhender la dispersion à longue distance du pathogène et évaluer quels type de paysage a le plus impacté l'invasion de la maladie.

Pour répondre à ces objectifs, deux axes de recherche sont proposés :

- Le premier est la modélisation de la maladie à l'échelle de la plante par le développement du modèle hôte-pathogène baptisé "**SiBaToka**". Ce modèle est basé sur la description mécaniste, d'une part, de la croissance de la plante sur plusieurs cycles de culture et d'autre part, du cycle épidémique du pathogène *Mycosphaerella fijiensis*. Ces descriptions ont été établies en conditions environnementales optimales, comme le sont celles des milieux tropicaux.
- Le deuxième est la modélisation de l'invasion de la cercosporiose noire à l'échelle régionale de la Martinique au travers d'un modèle stochastique basé sur des processus ponctuels spatio-temporels.

Chapitre 2

Modèle de simulation de la cercosporiose noire à l'échelle d'un bananier : le modèle SiBaToka

Sommaire

2.1 Article en préparation	24
2.2 Simulation et évaluation du modèle	70
2.2.1 Description de la plante	71
2.2.2 Densité d'inoculum primaire	73
2.2.3 Dynamique des inoculums au cours des cycles de culture	76
2.2.4 Effet de l'inoculum primaire sur la vitesse du cycle infectieux	78
2.2.5 Effet de la résistance du bananier sur la dynamique parasitaire	78
2.3 Discussion	79

Nous présentons dans ce chapitre la construction et l'évaluation d'un modèle spécifique décrivant la dynamique du pathogène *Mycosphaerella fijiensis* sur son hôte en croissance : le modèle **SiBaToka** "Simulation on **Banana** of **SigaToka** * disease".

* La vallée de Sigatoka est une vallée des îles Fidji où la cercosporiose noire a été décrite en 1963 pour la première fois.

Dans la première partie du chapitre, nous présentons l'article que nous soumettrons dans la revue *Ecological Modelling*. Cet article décrit dans un premier temps les étapes de la construction du modèle : la simulation de la croissance d'un bananier sur plusieurs cycles de culture et la dynamique du pathogène *Mycosphaerella fijiensis* sur son hôte. Dans un second temps, nous présentons la bibliographie et les données qui nous ont permis de renseigner les paramètres décrivant la dynamique du pathosystème. L'incertitude liée aux paramètres les plus influents du cycle infectieux nous a amenés à utiliser une méthode bayésienne de type MCMC afin d'inférer ces paramètres. Pour ce faire, nous avons mis en place une expérimentation de suivi de la surface malade à petite échelle spatiale (plusieurs feuilles de plusieurs plantes) pendant trois mois, en République Dominicaine. L'inférence a permis d'affiner nos connaissances sur ces paramètres influents sur le pourcentage de maladie observé sur une variété sensible de plantain de type "Macho por Hembra". La validation du modèle a été effectuée sur des données dynamiques acquises antérieurement sur une variété plantain de type "French Clair" à plus grande échelle spatiale (évaluation du pourcentage de nécroses sur une plante entière sur plusieurs cycles de culture). Le modèle simule efficacement l'infection d'un bananier par la cercosporiose noire et peut être vu comme un outil de meilleure compréhension du pathosystème car tous les paramètres d'entrée dépendent de la variété et les simulations, peu coûteuses en temps de calcul, peuvent produire des sorties d'intérêt multiples.

Le développement du modèle a été initié en utilisant le logiciel R [R Development Core Team, 2008] puis pour des raisons de temps de calcul, le modèle a entièrement été codé en C++ (langage orienté objet) et les calculs numériques ont été réalisés à l'aide du calculateur du C3I, de l'université des Antilles.

Dans la seconde partie du chapitre, nous proposons d'exploiter l'outil développé pour mieux comprendre, par simulation, le fonctionnement épidémique de la maladie à l'échelle de la plante. En particulier la dynamique de l'inoculum secondaire pour laquelle peu de connaissances précises existent dans la littérature.

2.1 Article en préparation

Development and calibration of a host-pathogen model: the case of Black Leaf Streak Disease of bananas

C. Landry^a, F. Bonnot^b, J. Vaillant^c, C. Abadie^{a,*}

^aUMR BGPI-CIRAD, Neufchateau, Capesterre B.E., Guadeloupe, FWI

^bUMR BGPI-CIRAD, Montpellier, France

^cUniversité des Antilles, Pointe à pitre, Guadeloupe, FWI

Abstract

Black Leaf Streak Disease (BLSD) is considered as the most destructive foliar disease of bananas and is caused by the fungal pathogen *Mycosphaerella fijensis*. To improve knowledge on the pathogen dynamics and on the components of partial resistance of varieties, a mechanistic simulation model of BLSD was designed, calibrated and validated. The model runs in discrete time at plant scale and describes the development of the plant and of the pathogen over several plant cycles under optimal epidemiological conditions. The model includes two sub-models: the plant sub-model is deterministic and simulates a simplified architecture and growth of the banana, and the epidemic sub-model simulates the detailed epidemic cycle including infection process, growth of the lesions on the leaves, asexual and sexual sporulation, and spore dispersal on the plant. A global analysis of sensitivity of the model allowed to identify the main influence of three epidemiological parameters: infection efficiency, incubation period, and lesion growth rate. These three parameters were estimated in a Bayesian framework using Markov Chain Monte Carlo methods with data of dynamics of lesions on the leaves collected under natural conditions. The model was validated at plant scale and for two successive plant cycles. The posterior values provide precise knowledge on the life traits for one variety. Simulations with

*Corresponding author

Email addresses: clara.landry@cirad.fr (C. Landry), francois.bonnot@cirad.fr (F. Bonnot), jean.vaillant@univ-ag.fr (J. Vaillant), catherine.abadie@cirad.fr (C. Abadie)

the inferred parameters allowed to conclude that the lesion growth rate and the plant growth rate impact the disease severity. The model developed here will be used as a tool to understand the BLSD epidemic and will be very helpful to design and evaluate efficient control practices.

Keywords: BLSD, foliar epidemic, host-pathogen model, Bayesian inference

1. Introduction

Host-pathogen modeling attempts to describe mathematically, or computationally, biological knowledge in order to better understand the functioning of host-pathogen interactions. Host-pathogen models are effective tools to develop decision support for the control of airborne fungal diseases of major field crops [33]. As "emerging infectious diseases caused by fungi are representing a world-wide threat to food security" [23] and given high dependency between the host dynamics and the transmission of pathogens ([17][22]), many mechanistic models describing interactions among host, pathogens, and environment have been proposed to analyze foliar epidemics [20]. The understanding of the spread of a plant pathogen in time and space requires a good knowledge on the components of the epidemic cycle that affect the disease progress and on spore dispersal gradients. The components of the epidemic cycle include infection efficiency, latent period, lesion size and growth, spore production rate and infectious period [35]. Spore dispersal gradients [56] concern the primary inoculum, which initiates the epidemics, and the secondary inoculum including the auto-infection processes (infection by spores produced on the same plant) that considerably increases the disease level on a plant scale [36].

At plant scale, mechanistic models have been developed on key plant pathogens, describing precisely the steps of the fungal disease epidemic cycle ([52] [11], [5], [64] [55], [53]). These mechanistic models describe the disease dynamic on a growing plant and include plant architecture knowledge either in the same model ([11], [49]) or using some plant growth specific model already developed ([43], [5]). Some generic epidemiological models have been developed and al-

low to analyse the epidemics at various scales (infection sites to plant) ([63], [28]). At crop scale, some generic airborne diseases models simulate precisely epidemic cycles ([13], [14]). To assess the effect of host resistance on disease epidemics, some epidemiological models have been developed at crop scale [39] or landscape scale [46].

To quantify the effect of auto-infection on the evolution of an epidemic cycle component [7] used the proportion of spores remaining on the parent leaf. [63] assessed the effect of spore deposition on epidemiological dynamics at three scales: auto-deposition (lesion scale), allo-leaf-deposition (leaf scale) and allo-plant-deposition (plant scale) assuming that spore dispersal occurs at one of the three levels according to three known ratios. [54] used directly the proportion of spores reaching the leaf surface. Dispersal function can also be used. [11] defined a dispersal function depending on a cone of angle determined by wind direction and distance between source and deposit place. [5] used both dispersal functions for ground and leaf inoculums, and a multiplication factor to take account of leaf self-infection. All these models include climatic factors because, in temperate regions, seasons highly influences the dynamics of the epidemic cycle and host development. In tropical conditions, climatic conditions are usually favorable plant disease development, avoiding a drastic extinction of plant pathogen populations like in temperate regions.

In tropical regions, one of the major plant diseases affecting bananas and plantains is the Black Leaf Streak Disease (BLSD) or Black Sigatoka, caused by the ascomycete fungus *Mycosphaerella fijiensis*. BLSD is considered as the most damaging and costly foliar disease for commercial banana and plantain production [34], because it causes yield losses from 20 to 80% in the absence of fungicides [15]. The use of chemical treatments is environmentally non durable and economically inappropriate for small producers [41]. Breeding for resistance appears to be the most appropriate way to sustainably control BLSD ([34] [6]). Banana resistance level may depend on aggressiveness (quantitative pathogenicity) of *Mycosphaerella fijiensis* [12] and on various plant resistance components acting on several steps of BLSD epidemic cycle ([1],[25] [32]). BLSD can dis-

perse at long distance [10]. [3] set up a field experiment and found dispersal events at distances of six kilometers from the source of infection.

Due to field and laboratory experiments, some steps of *Mycosphaerella fijiensis* epidemic cycle are well known: quantitative data are available on the leaf receptivity and on sporulation (duration and capacity) ([32], [24], [1]). This knowledge offers an opportunity to analyse the impact of the auto-inoculum process and host resistance on the epidemics.

To the best of our knowledge, no model of *Mycosphaerella fijiensis* population dynamics at plant scale has been yet developed, except the mention of the development of a mechanistic model on BLSD in an international congress [4]. Two banana crop models have been developed by [8] and [60], but they were based on leaf area index (LAI) and did not account for precise plant description in space.

Our objectives were to develop a specific mechanistic model that simulate the dynamics of the banana/*Mycosphaerella fijiensis* pathosystem over several plant cycle and to estimate the main epidemic cycle parameters using good prior knowledge and field experimental data, in order to improve the understanding on this pathosystem.

In this paper, we first present the general structure of the model as follows: the plant sub-model, which includes the phenological cycle, the plant architecture and growth, and the epidemic sub-model, divided in four main steps: infection, lesion growth, sporulation and spore dispersal. Second, we present the experimental data acquired for model calibration and the bayesian method used to estimate the main parameters of the epidemic cycle. Third we present the validation of the model using a time series data and we present contrasted simulation of BLSD dynamics to identify the impact of plant resistance components.

2. Materiel and methods

2.1. Biological background

85 Bananas and plantains are herbaceous flowering plants that produce successive pseudostems whose terminal bud produces the inflorescence and the bunch. A plant cycle comprises four main stages: planting (for the first cycle) or sucker emergence (lateral shoot for the following plants), growth, flowering, and harvest immediately followed by the elimination of the plant.

90 The first plant cycle lasts from planting to the elimination of the first harvested plant, called mother plant. The following cycles overlap so that each cycle lasts from the sucker emergence, which occurs during the growing period of the plant, until the elimination of the harvested plant derived from the sucker (Figure 1). The lengths of the four main stages depend on both cultural practices (land preparation, pruning, weed control, and fertilization) and climatic
95 conditions (temperature, radiation, and wind) [58]. The total plant cycle lasts about three hundred days.

During the growing period, the banana produces functional leaves, which are preformed inside the pseudostem, after which the flowering starts and the
100 leaf emergence stops [58]. The leaf emergence rate decreases during the growing period to become constant when the flower stalk starts to rise inside the pseudostem [37]. The leaf emergence rate is lower for the suckers than for the mother plant because the growth of each sucker is inhibited by the previous plant. The life duration of a banana leaf lasts between seven and ten weeks [59].

105 The banana leaf is composed of a petiole and two symmetrical blades, which are rolled-up inside the pseudostem before leaf emergence. The left blade unfolds first, followed by the right blade until the complete emergence of the leaf. The surface area of a leaf remains constant after emergence, but successive leaves are larger and larger [37].

110 The epidemic cycle of *Mycosphaerella fijiensis* is complex because of the occurrence of both asexual spores (conidia) and sexual spores (ascospores). The epidemic cycle lasts about three weeks (for a susceptible variety) and can be

summarized in four main steps: infection, lesion growth, sporulation, and spore dispersal [15].

115 Infection consists in the deposition of spores on the leaf surface and the penetration of germinated spores into the host foliar tissues through stomata. After an incubation period, very small lesions (specks) appear [41].

Lesions grow at constant logistic rate [21] and the asexual sporulation starts when lesions have reached a minimum size. During the asexual sporulation step, 120 lesions produce conidia in number proportional to lesion size and to sporulation capacity, which can disperse mainly by rain on the same leaf or older leaves or by wind at an average distance of approximately three meters [47]. Conidias can be considered as the major source of auto-inoculum at plant scale [9]. After the asexual sporulation step, the growing lesions become sexually mature. When 125 two lesions of opposite mating type coalesce, they become spot lesions and produce ascospores ([15], [29], [12]), which are mainly dispersed by wind and are considered as the most important source of inoculum for the long-distance dispersal with an average dispersal distance of approximately 360 m [47]. Conidias and ascospores are both infective, and germinate in the presence of high relative 130 humidity or free water [42], but receptivity of leaves to infection is known to decrease with leaf age [34]. Considering both the life duration of a banana leaf and the epidemic cycle duration, all lesion stages are typically present together on the plant.

According to field observations, lesions are distributed on the leaf following 135 a symptom density gradient [34]. The most infected part of the leaf is the top of the left blade and the least infected part is the bottom of the right blade. It is assumed that most infections occur on new leaves, between leaf emergence and complete unfurling, by spores transported by wind [58]. As the contamination by *Mycosphaerella fijiensis* is highly dependent of leaf emergence, it is essential 140 to study both the banana plant cycle and *Mycosphaerella fijiensis* epidemic cycle to better understand the BLSD dynamics.

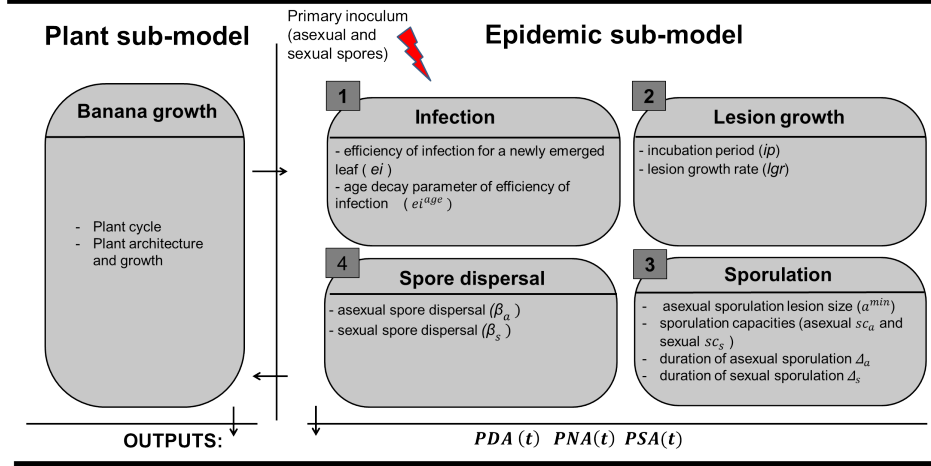


Figure 1: Overview of the host-pathogen model(banana/BLSD) general framework and the articulation of the plant sub-model and the epidemic sub-model. Only parameters of the epidemic sub-model are mentionned. PDA is the percentage of disease area, PNA is the percentage of necrotic area and PSA is the percentage of asexual sporulating area.

2.2. Model structure

To describe the BLSD dynamics at plant scale, we constructed a host-pathogen model using a mechanistic formalism in discrete-time.

As the main objective of the host-pathogen model is to study the effect of host resistance on disease level, all the model input parameters, including the epidemic parameters, are variety-dependent (Table 2). Model outputs are the disease quantitative indicators commonly used for field disease evaluation: PDA, the percentage of disease area and PNA the percentage of necrotic area [12]. The model runs at a daily time step and comprises two sub-models: the plant sub-model and the epidemic sub-model (Figure 1).

The plant sub-model is spatially explicit and is based on a simplified description of the architecture and the growth of banana under optimal agronomic and climatic conditions (Figure 2). The development of the successive plants is deterministically governed by a set of rules and equations depending on parameters issued from literature and field experiments. The objective of this sub-model

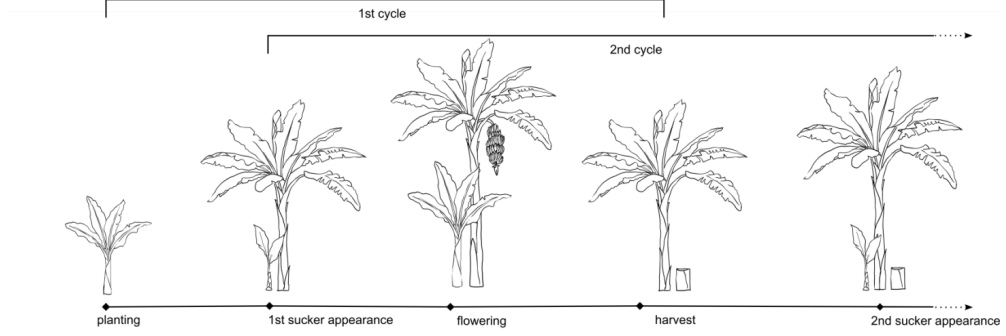


Figure 2: Plant cycle. The four successive development stages are planting, sucker appearance, flowering, and harvest followed by the elimination of the harvested plant.

is to provide values, at each time unit, of the areas of leaf elements (defined further in the text) and of the distances between pairs of leaf elements.

The epidemic sub-model is based on the four main steps of *Mycosphaerella*
 160 *fijiensis* epidemic cycle. We make the assumption that spore dispersal between
 any two leaf elements depends on the areas of the two elements and on their
 separation distance through a dispersal function. The two elements can belong
 to the same leaf or to different leaves, and the leaves can belong to the same
 plant or to different suckers.

165 The variables used in the two sub-models are listed in Table 1 and the input
 parameters in Table 2.

Sub-model	Symbol	Variable	Equation
Plant	θ	leaf azimuth	-
	ϕ	leaf inclination	-
	z	height of the leaf insertion point	-
	$r(t)$	leaf emergence rate at time t	-
	t_L^{init}	time of the leaf emergence	3
	L_L	length of blade	1
	W_L	width of blade	1
	A_L	area of the leaf	1
	p_i	proportion of open area in compartment i	-
	(r_1, \dots, r_7)	radial distances from compartments to the pseudostem	2
	$(\theta_1, \dots, \theta_7)$	centroid azimuths of compartments	2
	(z_1, \dots, z_7)	centroid heights of compartments	2
	d_{ij}	distance between two compartment centroids	2 9
Epidemic	E	infection efficiency	3
	A_0^j	open area in leaf compartment j	4 7 9 12 13
	A_f^j	lesion free area in leaf compartment j	4
	A^j	area of the leaf compartment j	4
	N^{in}	total number of spores deposited on a compartment	4 13
	N_l^{new}	number of new lesions	4
	N_l^i	number of lesions in a cohort i	7
	A_l^i	cohort i area	5 6
	N_a^{out}	number of asexual spores emitted by a cohort	6
	q_i	probability of meeting between two mature lesions	7
	$N_l^{co,i}$	number of coalesced lesions	-
	$A_l^{co,i}$	area of coalesced lesions	8
	N_s^{out}	number of sexual spores emitted by a cohort	9
	N_a^{in}	number of asexual spores received by a compartment	9 13
	N_s^{in}	number of sexual spores received by a compartment	12 13
	f_a	dispersal function of asexual spores	9
	f_s	dispersal function of sexual spores	-
	$g(\beta)$	empirical normalizing function of the dispersal function	10

Table 1: Variables used in the model

Sub-model	Symbol	Variable	Units	Plantain mean values	variety	Source
Plant	$N_L^{0,init}$	number of functional leaves on the mother plant at planting	leaves	3		acquired data
	$N_L^{1,init}$	number of leaves on the sucker at the harvest of the mother plant	leaves	10		[38]
	N_L^{final}	number of leaves issued before reaching the final leaf	leaves	18		[37]
	ler^{init}	emission rate				
		leaf emergence rate at planting(mother plant) or at harvest of the mother plant (sucker)	compart./day	1.32		[59]
	ler^{final}	leaf emergence rate at the end of the cycle	compart./day	1		[59]
	N_L^{max}	maximum number of leaves on the plant (average for mother plant and suckers)	leaves	12		[32]
	$N_L^{harvest}$	number of leaves on the plant at harvest	leaves	5		[32]
	N_L^{flower}	number of functional leaves emerged until flowering	leaves	40		[50]
	N_L^0	number of leaves on the mother plant at the sucker emergence	leaves	18		[37]
	$\Delta_f - h$	duration between flowering and harvest (average for mother plant and suckers)	days	90		[32]

(a_1, \dots, a_7)	proportion of surface area of each compartment	[0,1]	0.055, 0.205, 0.055, [27] 0.205, 0.27, 0.105, 0.105	
ϕ_L^V	inclination difference between two consecutive leaves	degree	156	[37]
ϕ_L^H	phyllotaxis angle	degree	10	[27]
H_L^0	height of the insertion point of the lowermost leaf at plant mother planting	mm	360	acquired data
H_L^1	height of the insertion point of the lowermost leaf at sucker emergence	mm	449	acquired data
l_{node}	length of internode	mm	87	acquired data
(l_x^1, \dots, l_x^7)	relative x coordinates of the compartment centers from the base of the blade	[0,1]	0.832, 0.495, 0.832, 0.495, 0.610, 0.148, 0.148	acquired data
(l_y^1, \dots, l_y^7)	relative y coordinates of the compartment centers from the base of the blade	[0,1]	0.541, 0.628, -0.541, -0.628, 0.377, -0.377	acquired data
L_L^{max}	blade length of the largest leaf	mm	2016	acquired data
L_L^{min}	blade length of the smallest leaf	mm	486	acquired data
W_L^{max}	blade width of the largest leaf	mm	846	acquired data
W_L^{min}	blade width of the smallest leaf	mm	253	acquired data
l_p^0	ratio petiole length / blade length (mother plant)	[0,1]	0.162	acquired data
l_p^1	ratio petiole length / blade length (sucker)	[0,1]	0.238	acquired data
n_a	initial asexual inoculum	spores/mm ²	-	-

n_s	initial sexual inoculum	spores/mm ²	-	-
ei	efficiency of the infection for a newly emerged leaf	number of lesions/spore	[0.0004, 0.01853]*	acquired data
ip	incubation period : time between infection and leaf spot appearance	days	[11, 36]	[24]
lgr	lesion growth rate	mm ² /days	[0.01, 0.33]*	acquired data
ef^{age}	age effect on infection efficiency	leaf/days	0.08	-
a^{min}	lower bound of lesion size releasing asexual spores	mm ²	2	[24]
sc^a	asexual sporulation capacity: number of asexual spores released	number/mm ²	7	[24]
sc^s	sexual sporulation capacity: number of sexual spores released	number/mm ²	182	[26]
Δ_{a-s}	minimum duration between asexual and sexual sporulation periods	days	10	Abadie, pers.data
Δ_a	asexual sporulation duration	days	8	acquired data
Δ_s	sexual sporulation duration	days	1	[34]
β_a	decay parameter of the asexual dispersal function	mm ⁻¹	0.001	[47]
β_s	decay parameter of the sexual dispersal function	mm ⁻¹	0.00001	[47]

Table 2: Input model parameters used in SiBaToka. The interval values

noted * are the prior interval used for model calibration.

2.2.1. Plant sub-model

Plant cycle. The mother plant, grown under greenhouse conditions in disease free environment before planting, is initialized with $N_L^{0,init}$ leaves, and each
 170 sucker appears with one functional leaf when the previous plant bears N_L^0 leaves.

During the growing period, N_L^{flower} leaves are produced from the emergence of the first leaf until flowering. The oldest leaves are regularly removed so that the number of leaves on the plant cannot exceed N_L^{max} . After flowering and during a period of length Δ_{f-h} necessary for bunch development, leaf emergence
 175 stops and the number of leaves on the plant decreases linearly with time from N_L^{max} to $N_L^{harvest}$ leaves at harvest.

The leaf emergence rate decreases linearly with time from ler^{init} at planting (for the mother plant) or at harvest of the previous plant (for the sucker) to ler^{final} when N_L^{final} leaves have emerged, then remains constant and equal to
 180 ler^{final} until flowering. The leaf emergence rate of the sucker before the harvest of the previous plant is taken as constant and calculated to get $N_L^{1,init}$ leaves emerged from the sucker at the harvest of the previous plant.

Plant architecture and growth. Each plant is modeled as a series of leaves inserted on a vertical pseudostem. The position of each leaf in space is charac-
 185 terized by three values: the height of the insertion point on the pseudostem, the leaf azimuth (angle from north in a clockwise direction) and the leaf inclination (angle from vertical). The height z of the insertion point is determined by the insertion height of the first leaf (H_L^0 for the mother plant and $H_L^1 > H_L^0$ for each successive sucker, which are taller than the mother plant) and by the
 190 length L_{node} of the internode between two successive leaves.

The leaf azimuth θ is determined by the phyllotaxis angle ϕ_L^H between two successive leaves, the azimuth of the first leaf being arbitrarily fixed to 0° . Each leaf emerges vertically and its inclination increases at every time step by the quantity $r(t)\phi_L^V$, $r(t)$ being the leaf emergence rate at time t and ϕ_L^V the
 195 inclination difference between two consecutive leaves, taken as constant. The successive suckers are supposed to be at the same spatial location as the mother

plant.

Leaf is modeled as two symmetric blades and a petiole whose length is a proportion $l_p = l_p^0$ (for the mother plant) or l_p^1 (for suckers) of the blade length. In our simple architectural model, all the parts of the leaf are in a same plane. The contour of the largest leaf (the last issued) is determined by a set of 20 point coordinates measured on a blade, and the contour of the other leaves is obtained by a linear transformation of these coordinates given the length and the width of the blade. Length L_L and width W_L of the N^{th} blade are functions of N that fit the observed data according to the curves and the foliar area approximation given by [37].

$$L_L = L_L^{min} + \frac{N - 1}{N_L^{flower} - 1} (L_L^{max} - L_L^{min})$$

$$W_L = W_L^{max} - (W_L^{max} - W_L^{min}) \exp(-0.05(N - 1)) \quad (1)$$

$$A_L = 0.8 L_L W_L$$

In order to model the opening process of the leaf, we divided the foliar surface into seven compartments corresponding to the opening order of the foliar surface, compartment 1 being the first open and compartment 7 the last open. Each compartment c is attributed a proportion p_c of open area. When the leaf is rolled, $p_c = 0 \forall c$. When the leaf opens, $p_1, \dots, p_k (k \leq 7)$ are successively set to 1 excepted for p_k , which can remain lower than 1 if compartment k is not completely open. Opening new compartments (whose number can be integer or not) is realized by filling the compartments of the last open leaf and if necessary by making new leaves to emerge with the remaining number of compartments. For each compartment c , the proportion a_c of foliar area ($\sum_1^7 a_c = 1$) and the coordinates (l_x^c, l_y^c) of the centroid relative to the base of the leaf blade, expressed as proportions of leaf blade length (for l_x^c) and width (for l_y^c), were determined from the contour coordinates of the compartment. The absolute coordinates x and y of the centroid relatively to the petiole extremity of a leaf blade of length L_L and width W_L are given by $x = L_L(l_p + l_x^c)$ and $y = W_L l_y^c / 2$. Given the

azimuth θ and the inclination ϕ of the leaf and the height z of the insertion point, the spatial coordinates (r_c, θ_c, z_c) of centroid c , where r_c is the radial distance from the pseudostem, θ_c the centroid azimuth, and z_c the centroid height, are given by equations

$$r_c = \sqrt{(x \sin \phi)^2 + y^2}$$

$$\theta_c = \theta + \tan^{-1}(y/(x \sin \phi))$$

$$z_c = z + x \cos \phi$$

If i and j are any two compartment centroids of coordinates (r_i, θ_i, z_i) and (r_j, θ_j, z_j) , their separating distance is given by

$$d_{ij} = \sqrt{(r_i^2 + r_j^2) - 2r_i r_j \cos(\theta_i - \theta_j) + (z_i - z_j)^2} \quad (2)$$

2.2.2. Epidemic sub-model

210 *Infection.* In the epidemic sub-model, the infection step corresponds to the daily creation of new lesions from deposited spores on each leaf compartment. The foliar surface is considered as a set of contiguous sites where spores can be deposited and germinate. Each site corresponds to the surface that can be occupied by a new lesion. The site area was fixed to 1 unit area (1 mm²) in
 215 the model, based on lesion observations. Only a proportion E of spores, corresponding to the infection efficiency, effectively germinate if they are deposited on vacant sites. The infection efficiency E is assumed to be the same for asexual and sexual spores [34] and to decrease exponentially with leaf age to model the lower receptivity of older leaves to infection [59]:

$$E = ei \cdot \exp(-ei^{age}(t - t_L^{init})) \quad (3)$$

220 where ei is a parameter to be estimated corresponding to the spore infection efficiency on a newly emerged leaf, ei^{age} is the rate of decrease of infection efficiency with leaf age and t_L^{init} is the time of leaf emergence.

For a small quantity ΔX of spores deposited on a leaf compartment j with open area A_0^j and lesion-free area $A_f^j \leq A_0^j$, the quantity of new lesions is $E A_f^j / A_0^j \Delta X$. Integrating over the total number of spores deposited N^{in} gives the number of new lesions

$$N_l^{new} = A_f^j \left[1 - \exp \left(-E \frac{N^{in}}{A_0^j} \right) \right] \quad (4)$$

The exponential function appears because of the saturation caused by the possible deposition of several spores on the same site.

In the model, the new lesions created in each compartment at each infection time constitute a cohort i of lesions for which the number N_l^i and the total area A_l^i are updated at each time step.

Lesion growth. The new cohort i contains N_l^{new} lesions with total area $A_l^i = N_l^{new}$ since each lesion size is 1 unit area. The area A_l^i remains constant during the incubation period of length ip (parameter to be estimated), then increases according to a logistic growth function:

$$A_l^i(t) = lgr A_l^i(t-1) \left(1 - \frac{A_l^i(t-1)}{\sum A_l^i(t-1)} \right) \quad (5)$$

where lgr (a parameter to be estimated) is the lesion growth rate and the summation is over all the cohorts in the compartment. Lesion growth continues at the same rate during asexual and sexual sporulation.

Sporulation. The asexual sporulation for a cohort i of lesions starts when the individual lesion area A_l^i / N_l^i reaches the value a^{min} and lasts until the end of the sporulation period of length Δ_a . During this period, as each lesion produces sc_a spores per time and area unit, the number of asexual spores $N_a^{out}(t)$ emitted by the cohort at time t is:

$$N_a^{out}(t) = sc_a A_l^i(t) \quad (6)$$

Lesions become sexually mature after a period of length Δ_{a-s} that starts at the end of asexual sporulation. When at least two mature lesions of opposite mating type coalesce, they start emitting sexual spores. The probability

that a lesion of mature cohort i meets another lesion in any mature cohort in compartment k is modeled as

$$q_i = 1 - \left(1 - \frac{4s_i}{A_0^k}\right)^{N_i^i-1} \prod_{j, j \neq i} \left(1 - \frac{(\sqrt{s_i} + \sqrt{s_j})^2}{A_0^k}\right)^{N_l^j} \quad (7)$$

where the product is over mature cohorts, $s_i = A_l^i/N_l^i$, A_l^i and N_l^i being the total area and the number of lesions in cohort i , and A_0^k is the open area in compartment k (cf. Appendix 6).

At each time step and for each mature cohort i , the number of coalescent lesions $N_l^{co,i}$ that meet a sexually compatible lesion in any mature cohort is drawn in a binomial distribution with size N_l^i and probability $1/2 q_i$. Then the number $N_l^{co,i}$ and the corresponding area $A_l^{co,i} = N_l^{co,i} \cdot (A_l^i/N_l^i)$ of coalesced lesions are subtracted from each mature cohort i , where only mature lesions that have not coalesced remain. The new cohort contains $\sum_i N_l^{co,i}$ coalesced lesions sexually sporulating of surface area $\sum_i A_l^{co,i}$. Sexual sporulation stops after a sporulation period of length Δ_s . During this period, the cohort continues to grow and emits at each time unit t a number of sexual spores $N_s^{out}(t)$:

$$N_s^{out}(t) = sc_s \sum_i A_l^{co,i} \quad (8)$$

where sc_s is the sexual sporulation capacity and the sum is over all the coalescent cohorts area in the compartment. At the end of sexual sporulation, the cohort stops growing but remains in the compartment as a necrotic surface. The number of sexual (resp. asexual) spores emitted by a compartment at each time step is the sum of the numbers of sexual (resp. asexual) spores emitted by all the cohorts in the compartment.

Spore dispersal. Let $N_a^{out}(i, t)$ be the number of asexual spores emitted by compartment i at time t . We model the number of asexual spores received by compartment j from compartment i at time t by

$$N_a^{in}(i, j, t) = N_a^{out}(i, t) \cdot f_a(d_{ij}) \cdot A_0^j(t) \quad (9)$$

where f_a is the dispersal function of asexual spores, d_{ij} is the distance between the centroids of compartments i and j , and A_0^j is the open area of compartment j . If $i = j$, $d_{ij} = 0$ and $N_a^{in}(i, i, t)$ is the asexual auto-inoculum of compartment i at time t .

The total number of asexual spores received at time t by compartment j is therefore $N_a^{in}(j, t) = \sum_i N_a^{in}(i, j, t)$, where the summation is over all the compartments of the same plant and of the plant of other cycle existing at time t . We used an exponential dispersal function $f_a(d_{ij}) = g(\beta_a) \exp(-\beta_a d_{ij})$, where β_a is the distance decay parameter and $g(\beta_a)$ is an empirical normalizing function such that $\forall i, \sum_j f_a(d_{ij}) A^j = g(\beta_a) \sum_j \exp(-\beta_a d_{ij}) A^j \leq 1$. We used

$$g(\beta_a) = \frac{1}{\max_i (\sum_j \exp(-\beta_a d_{ij}) A^j)} \quad (10)$$

where A^j is the area of the compartment j and the maximum is taken over all the compartments of all the architectural configurations occurring during plant growth. This function was well fitted by a nonlinear model

$$\ln[g(\beta_a)] = \ln(g_{\min}) + [\ln(g_{\max}) - \ln(g_{\min})](1 - \exp(-c\beta_a)) \quad (11)$$

where g_{\min} , g_{\max} , a , and c are regression coefficients.

The dispersal of sexual spores is modeled similarly, excepted that the distance decay parameter β_s of the exponential dispersal function f_s differs from β_a . The normalizing function $g(\beta_s)$ is the same as for asexual spores since it depends only on the plant architecture. The equations of the sexual spore dispersal, with trivial notation, become

$$N_s^{in}(i, j, t) = N_s^{out}(i, t) \cdot f_s(d_{ij}) \cdot A_0^j(t) \quad (12)$$

and $N_s^{in}(j, t) = \sum_i N_s^{in}(i, j, t)$.

The number of spores $n_p(t) \cdot A_0^j(t)$ necessary to initiate the epidemic cycle (primary inoculum) is added to each compartment j at time t . The quantity $n_p(t)$, which can be constant over the simulation or only during selected steps, corresponds to the number of asexual and sexual spores in the air that are deposited on the leaves per area unit and time unit.

Finally, the total number of asexual and sexual spores received at time t by compartment j is the summation of primary inoculum and the secondary inoculum corresponding to the spores produced during the asexual and sexual sporulation periods calculated as

$$N^{in} = n_p(t) A_0^j(t) + N_a^{in}(j, t) + N_s^{in}(j, t) \quad (13)$$

2.2.3. Model outputs

At each time t , the model can output a series of values corresponding to the state of the system for the plant sub-model and the epidemic sub-model. The values can concern an elementary part of the system (cohort of lesions or leaf compartment) or can be aggregated by leaf or plant. In the following, we focus on three quantitative indicators of the disease aggregated at the plant level: the percentage of diseased area (*PDA*), the percentage of necrotic area (*PNA*), and the percentage of asexual sporulating area (*PSA*), defined for plant p as:

$$PDA^p(t) = \sum_i A_l^i(t) / \sum_j A_0^j(t) \times 100 \quad (14)$$

$$PNA^p(t) = \sum_i A_l^{co,i}(t) / \sum_j A_0^j(t) \times 100 \quad (15)$$

$$PSA^p(t) = \sum_i A_l^{a,i}(t) / \sum_j A_0^j(t) \times 100 \quad (16)$$

where $A_l^i(t)$, $A_l^{co,i}(t)$, and $A_l^{a,i}(t)$ are respectively the surface area of lesions, coalesced lesions, and asexually sporulating lesions in cohort i at time t , and the summations in the numerators are over all the cohorts in plant p . The quantity $\sum_j A_0^j(t)$ where $A_0^j(t)$ is the open leaf area in compartment j at time t and the summation is over all compartments in plant p is the open leaf area in plant p at time t .

2.3. Model calibration

We performed a preliminary sensitivity analysis to quantify the impact of the epidemic sub-model parameters (fixing the plant sub-model parameters)

and their interaction in the decomposition of the *PDA* variance. We used e-FAST global sensitivity analysis (methodology is given in [57]) with 55000 model evaluations. We chose to retain in the calibration the parameters that accounted for more than 10% of *PDA* variance: the lesion growth rate lgr , the efficiency of the infection for a newly emerged leaf ei and the incubation period ip . The other parameters were set at the value reported in the literature (Table 2) or field experiment described below.

2.3.1. Fixed parameters and prior values

Parameter values were based on both literature and data from experiments (Table 2).

Data for the plant sub-model. We realized measures on the architecture of plantain in a private plot in Guadeloupe (FWI). The plot (16.12056 N, -61.58806 W) was planted with 0.2 ha of a plantain cultivar (French Sombre) susceptible to BLSD. We chose randomly 10 plants at flowering. The disease was present on all the plants and suckers. On each plant, we measured the blade length and width of the largest leaf (resp. L_L^{max} and W_L^{max}). On each sucker, we measured the height of the insertion point of the lowermost leaf H_L^1 and we measured the lengths of the petiole and of the blade to calculate the petiole relative length l_p^1 . We realized measures at planting in a plantain plot in Guadeloupe. We chose randomly 10 plants and we measured the blade length and width of the largest leaf (L_L^{min} and W_L^{min}), the height of the insertion point of the lowermost leaf H_L^0 and the lengths of the petiole and of the blade to calculate the petiole relative length l_p^0 . We also counted the number of leaves present on each plant $N_L^{0,init}$. All these data were averaged over the 10 replicates (Table 2).

Data for the epidemic sub-model. We used two sets of data to estimate epidemiological parameters.

The first set was collected in 1998 under field conditions in an experimental plot at Njomb in Cameroon where six varieties of banana and one variety of plantain had been planted with 12 to 13 plants per variety under high inoculum

305 pressure. Measures were realized on all the plants of plantain and of the two varieties Pisang madu and Pisang berlin.

Five months after planting, the kinetic of the lesion growth was measured on five isolated new lesions on the same leaf of each of the plants of the three varieties by drawing, twice a week during three weeks, the outline of each lesion
310 on transparent squares of 25 cm². The lesion area for each date was estimated by using an image analyzer and we calculated the daily growth rate lgr of each lesion. The duration of asexual sporulation Δ_a of each lesion was estimated from the surface of the lesion, knowing the average size of sporulating lesions according to the description of [24].

315 The second set of data was collected in 2012 under controlled conditions. The efficiency of the infection for a newly emerged leaf ei was estimated on banana leaf pieces cultivated in greenhouse. The inoculations were performed by spraying 1 ml of a conidial suspension concentrated at 7500/ml on 25 cm² leaf fragments maintained on a survival medium according to [2]. The inoculated
320 leaf pieces were scanned 25 and 32 days after inoculation. The average number of lesions was determined for each leaf piece by using the image analyzer Visilog (version 7). The infection efficiency ei was the ratio between the number of lesions and the number of spores deposited on the fragment.

The range values obtained for lgr and ei were considered as prior values in
325 the calibration method (Table 2).

2.3.2. Model calibration

Data used for calibration. We set up an experiment in the region of Moca (Dominican Republic) (19.33822 N, 70.46910 W) from September to December 2012. The climatic conditions (the monthly average of rainfall was 156 mm) were fa-
330 vorable to banana and BLSD developments. The study site was a fungicide unsprayed private plot of 0.7 hectare of plantain (cultivar Macho por hembra) planted at a density of 1666 plants/ha designed with 12 rows of 20 bananas distant of 2 meters and infected by BLSD. The plot was managed according to local agronomical practices in order to minimize mineral and biotic constraints other

335 than BLSD. Three plants were randomly chosen (excluding the plot borders).
 For each plant and for each of three successive emerged leaves, 21 squares of 6
 \times 6 cm were randomly chosen on the lower leaf surface. The number of squares
 in each compartment was proportional to the area of the compartment. Visible
 lesions (specks, streaks and spots) were drawn in each square on transparent
 340 paper twice a week since their appearance to senescence, during approximately
 14 weeks. Transparent papers were scanned twice a week, and the area of le-
 sions were determined using the image analyzer software VISILOG (version 7).
 The proportion of area covered by symptoms was averaged over the squares
 in each compartment and reported as BLSD rate on the entire compartment.
 345 Altogether, 7 compartments \times 3 leaves \times 3 plants \times 21 times of PDA were used
 for parameter estimation.

Calibration method. We used a Bayesian framework to estimate the three epi-
 demiological parameters lgr , ei and di simultaneously, given the following sta-
 tistical model

$$Y_{\ell,c}^p(t) = y_{\ell,c}^p(t) + \epsilon_{\ell,c}^p(t) \quad \text{with} \quad \epsilon_{\ell,c}^p(t) \sim N(0, \sigma^2) \quad (17)$$

with

$$\begin{aligned}
 Y_{\ell,c}^p(t) &= \sin^{-1}(\sqrt{PDA_{\ell,c}^p(t)/100}) \\
 y_{\ell,c}^p(t) &= \sin^{-1}(\sqrt{\widehat{PDA}_{\ell,c}^p(t)/100})
 \end{aligned}$$

where $Y_{\ell,c}^p(t)$ correspond to the t^{th} PDA observation in the compartment c of
 leaf ℓ on the plant p and $y_{\ell,c}^p(t)$ the correspondent to the \widehat{PDA} value predicted
 by the model conditionally to the values of all the model parameters and the
 350 known leaf emission rate of each plant.

The errors of the model prediction are assumed independent with residual
 variance σ^2 considered as an additional unknown parameter.

The Bayesian framework use prior information to estimate parameters val-
 ues by their posterior distribution. By definition, the posterior distribution is
 proportional to the product of the likelihood of the model and the parameter

prior distribution. Let $\theta = \{lgr, ei, di\}$ the set of parameters to be estimated and $Y^p = \{Y_{1,1}^p(1), \dots, Y_{3,7}^p(21)\}$ the data measured on each plant $p = \{1, 2, 3\}$, the posterior distribution of θ is

$$\pi(\theta|Y) \propto \pi(Y|\theta)\pi(\theta) \quad (18)$$

with $\pi(\theta) = \pi(lgr)\pi(ei)\pi(di)$ the priors distribution and $\pi(Y|\theta)$ the model likelihood defined as follow:

$$\pi(Y|\theta, \sigma^2) = \prod_{p=1}^3 \prod_{t=1}^{21} \prod_{\ell=1}^3 \prod_{c=1}^7 \frac{\exp[-(y_{\ell,c}^p(t) - Y_{\ell,c}^p(t))^2 / (2\sigma^2)]}{\sqrt{2\pi\sigma^2}} \quad (19)$$

The posterior densities of the four parameters were estimated using a Monte Carlo Markov Chain algorithm, more precisely using a Metropolis-Hastings algorithm within Gibbs algorithm (see [48] for details) described in Appendix 7 and used for example in [61], [31] or [45]. For the epidemiological parameters, we used uniform prior distribution on range described in Table 2 and for the residual variance, we used a non informativ Jeffrey prior $\pi(\sigma^2) \propto 1/\sigma^2$. The algorithm was validate on simulated data with different parameter values. The algorithm was run for 70000 iterations in three times (to calculate the convergence criterion of Gelman-Rubin [30]), the burn in period was chosen to 1000 iterations and a thinning of 5 iterations was used to remove auto-correlation in the chains. We used the mean of the posterior densities to estimate the parameters.

2.4. Model validation

Data used for validation. The model was validated on data of temporal kinetics of necrotic surface per plant (plantain) which were collected at Njombé (4.34°60'N, 9.39°0'E), in Cameroon by [16]. Five mini-plots of 28 plants (plantain) were planted and BLSD severity (percentage of necrotic surface per plant) was measured twice a month on 10 plants during the vegetative phase, at flowering and at harvest during the two plant cycles. A total of 650 data (13 dates \times 10 plants \times 5 replicates) was used for validation. The variance and mean for each date was calculated. Due to a lack of irrigation on the experimental plot

at planting, a delay of plant growth was observed with no leaf emergence for
 375 one month. The date of the first leaf emergence was estimated 28 days after
 planting for the simulation.

Method used for validation. Prediction quality was assessed by computing the
 root mean square error of the predictions (RMSE) as follows

$$RMSE = \sqrt{\frac{1}{N} \sum_{p=1}^2 \sum_{t=1}^{T_p} \left(PSA_p(t) - \widehat{PSA}_p(t) \right)^2} \quad (20)$$

where $\widehat{PSA}_p(t)$ is the model prediction of PSA for the plant p at time t condi-
 tionally to the mean posterior value of estimated parameters, N is the number
 of observations and T_p is the number of observations of each plant p . The RMSE
 380 has the same unit as PSA [62].

3. Results

3.1. Inference parameters

In the MCMC algorithm, the model simulations were run using primary in-
 oculum $n_a + n_s = 0.2$ spores/mm²/day. The Gelman-Rubin criterium $\hat{R} = 1$,
 385 calculated on the three chains replicates, indicates that we can assume con-
 vergence to the posterior distribution. The posterior distribution values are
 presented in table 3 and an example of the Markov chain and posterior density
 is available for the parameter lgr in Appendix 8.

Parameter	Posterior mean	Posterior SD	Q2.5%	Q97.5%
lgr	0.1674	2.034e-03	0.1641	0.1716
ei	0.0167	8.790e-04	0.01453	0.01832
di	31.013	0.42	30.12	32.04
σ^2	0.0664	3.194e-03	0.06042	0.07293

Table 3: Posterior mean, SD and Quantiles for each parameter inferred

Prior knowledge about these parameters were taken as large uniform distri-
 390 bution and posterior standard deviation are very small. The dynamical data

from plantain Macho por Hembra cultivar allow us to reduce the uncertainty in these parameters.

The posterior mean of the residual error variance (Eq. 17) is estimated as $\sigma^2 = 0.0664$. This refers to the arcsine transformed severity. For an average PDA of 50%, the standard deviation is about 26 % of PDA. This residual error takes account of both model error and measurement error in each compartment.

3.2. Model validation

The model simulations were run on two plant cycles (660 days) using primary inoculum $n_a + n_s = 0.4$ spores/mm²/day and fixed parameters defined in Table 2 and mean values of the posterior density of inferred parameters lgr , ei and di (Figure 3).

In this context, RMSE = 9.69%. The model reproduced the general trend for PNA observed in validation data.

The simulated curves increase within time with the highest level of disease level at harvest (this level increasing continuously from flowering). The disease dynamics is quite different between the two cycles. The PNA increases from 0% at planting to 85% at harvest when the PNA is at 55% on the sucker (beginning of the second cycle). The disease level decreases on the sucker from 55% to about 25% due to rapid leaf emission occurring after the elimination of the harvested mother plant. The disease level of the second cycle remains stable at about 25% until the flowering.

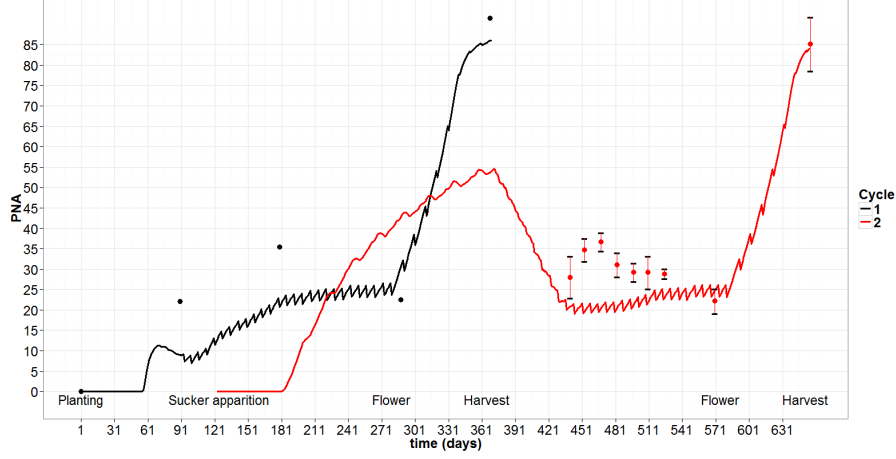


Figure 3: Dynamic of PNA on two consecutive plant cycles. Predicted PNA are represented by lines, and observed data by points, bounds around points in the second cycle represent standard deviation data.

3.3. Simulations

Effect of host resistance components. To illustrate the effect of host resistance components on the disease development, we chose contrasted values of realistic epidemic parameters for the three most influent parameters. The epidemic parameters were the inferred parameter $lgr = 0.1674$, $ei = 0.0167$, and $di = 31$ for plant A (plantain reference variety), $lgr = 0.092$, $ei = 0.0168$, and $di = 18$ for plant B (Pisang madu-like variety), and $lgr = 0.139$, $ei = 0.0197$ and $di = 18$ for plant C (Pisang berlin-like variety).

The model simulations were run on two plant cycles (660 days) using primary inoculum $n_a + n_s = 0.4$ spores/mm²/day. The values of the fixed parameters of the plant-submodel for the three varieties were the values indicated for plantain in Table 2. Figure 4 shows differences in PNA values between the three plants with the lowest PNA values for plant B and the highest for plant C. This difference occurs mainly before flowering.

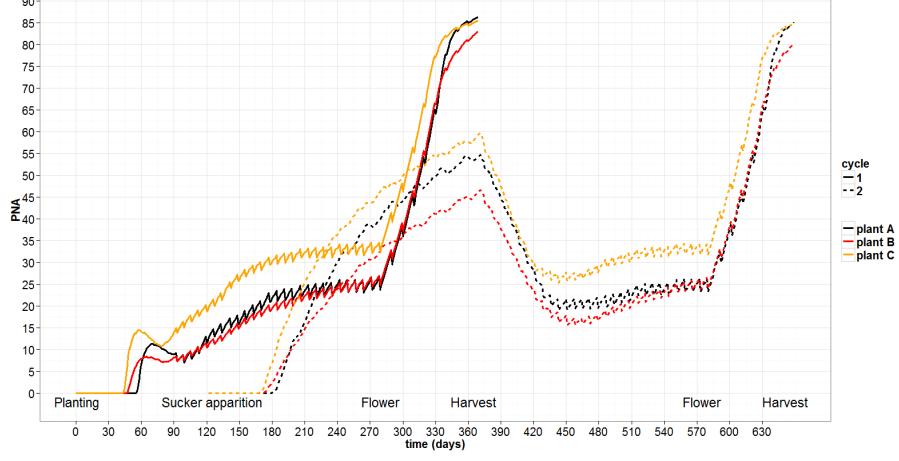


Figure 4: Dynamics of PNA per plant on the first and second plant cycles for plants A, B and C defined with different resistance component parameter values.

Effect of asexual spores dispersal. To illustrate the effect of asexual spore dispersal at a plant scale, we chose contrasted values of the decay parameter β_a of the dispersal function of asexual spores. Parameter simulations were defined with the mean of the inferred parameter values for lgr , di and ei (Table 3) and
430 with fixed parameters given in Table 2. For Plant A we chose $\beta_a = 0.0001$ i.e. parameter allowing long distance dispersal events and for Plant B we chose $\beta_a = 0.01$ which allowed only short distance dispersal events.

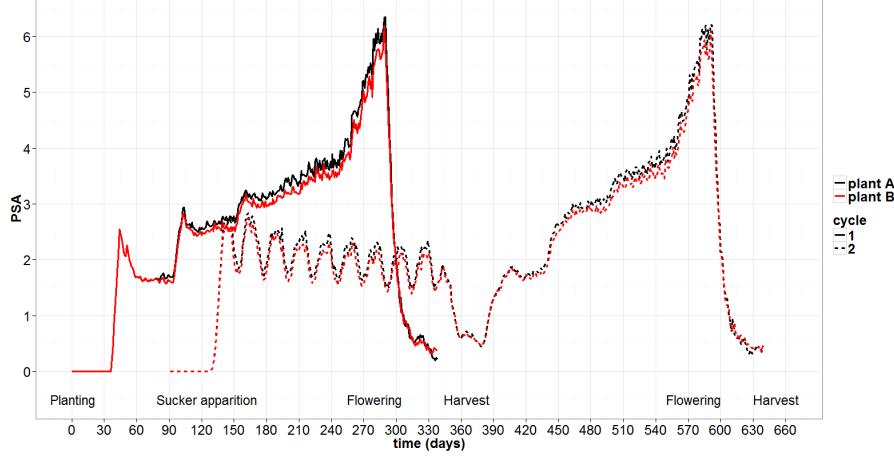


Figure 5: Dynamics of PNA per plant on two consecutive cycles for plants defined with different auto-inoculum parameter values : plant A ($\beta_a = 0,0001$ i.e. long distance dispersal), plant B ($\beta_a = 0.01$ i.e. short distance dispersal).

The model simulations were run on two plant cycles (660 days) using primary inoculum $n_a + n_s = 0.4$ spores/mm²/day. Figure 5 shows few differences of
435 PSA between plant A and B, the plant A with long distance dispersal has a slight higher percentage of asexual sporulating area than the plant B with short distance dispersal.

Effect of plant growth. To illustrate the effect of plant growth velocity on two
plant cycles, we chose contrasted values of the leaf emission rate ler^{final} . Pa-
440 rameter simulations were defined with the mean of the inferred parameter values for lgr , di and ei (Table 3) and with fixed parameters given in Table 2. For Plant A we chose $ler^{final} = 0.7$ i.e. a slow leaf emission rate and for Plant B we chose $ler^{final} = 1.33$ a rapid leaf emission rate.

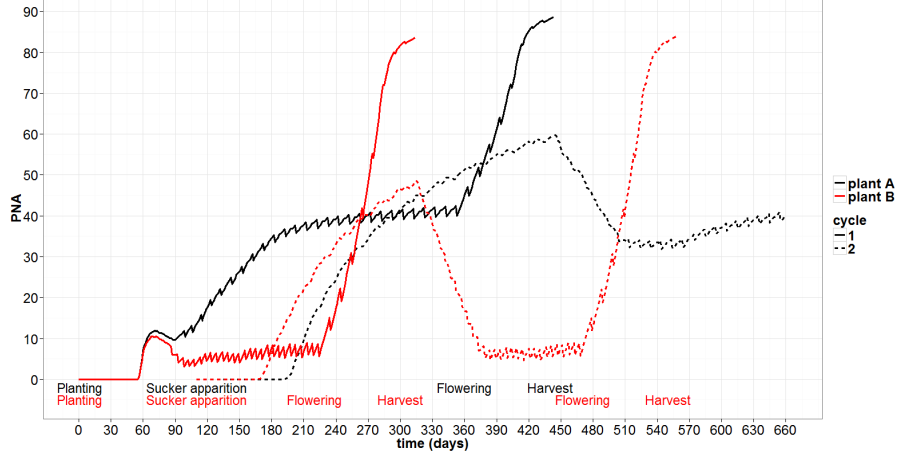


Figure 6: Dynamics of PNA per plant on 660 days for plants defined with different leaf emission rate values : plant A ($ler^{final} = 0.7$ i.e. , slow leaf emission rate), plant B ($ler^{final} = 1.33$, rapid leaf emission rate).

The model simulations were run on 660 days using primary inoculum $n_a +$
 445 $n_s = 0.4$ spores/mm²/day. Figure 6 shows high differences between the PDA
 of plant A and B. The percentage of necrotic area before flowering is about 8%
 for the plant B and 40% for the plant A. The plant B, with rapid leaf emission,
 takes about 220 days to flowering whereas the plant A takes 360 days.

4. Discussion

Our objectives were to develop a specific mechanistic model that simulates
 450 the dynamics of the banana/*Mycosphaerella fijiensis* pathosystem over several
 plant cycles and to estimate the main epidemic parameters using prior knowl-
 edge and field experimental data collected in Dominican Republic at leaf scale
 on a susceptible plantain variety, in order to better understand the epidemic
 455 behavior. We then validated the model on an independent dynamic dataset
 at plant scale from Cameroon. The percentage of necrotic area was accurately
 predicted over two plant cycles with a RMSE of 9.69%.

On banana, some simulation models at crop scale have been developed to study soilborne pests [60]: SiBaToka is the first model developed on BLSD of bananas, and our study is one of the first that estimates epidemiological parameters in a specific host-pathogen simulation model. It can be considered as a tool to better understand epidemics at plant scale, allowing to evaluate the effect of host resistance components or the effect of auto-infection. It can also be considered as a tool to evaluate the efficiency of control practices such as the effect of the host development on disease epidemics. Indeed, the model SiBaToka includes two sub-models that use variety-dependent input parameters: the plant sub-model, which simulates the plant cycle, the plant architecture and the plant growth, and the epidemic sub-model, which simulates the stages of the BLSD epidemic cycle (infection, lesion growth, sporulation and spore dispersal). Even if the impact of plant architectural traits on disease progress and disease level has been recently identified as a key component of disease control practices [17], it is not so frequent to find in the literature epidemiological models simulating both plant growth and plant disease dynamics. Recently, a mathematical model has showed analytically the effect of crop growth and canopy filtration on plant disease epidemics [22]. More recently, some generic models have been developed to simulate airborne fungal diseases by coupling plant or crop architecture and disease dynamics ([13] [14]) and by a functional-structural plant models [28]. Only few simulation models have modeled plant growth and specific airborne diseases as [11] did on the powdery mildew of grapevine, Robert 2008 on Septoria diseases of wheat or [40] on phoma-stem canker of oilseed. All these models use fixed parameters acquired from literature and experimental data.

The second aim of our study was to estimate the main epidemic parameters influencing the disease severity. Through a global analysis of sensitivity, three of the eleven parameters of the epidemic sub-model were identified to be accountable of more than ten percent on the percentage of disease area variance: the lesion growth rate lgr , the incubation duration di and the efficiency of the infection on a newly emerged leaf ei . The epidemiological effect of lesion growth on BLSD has been suggested through an experimental approach by [1]. Its im-

portance has already been demonstrated with the generic model of [13]. We
490 used a Bayesian framework based on Markov Chain Monte Carlo methods to
estimate the three parameters by taking account of available prior information.
A Metropolis-Hastings within Gibbs algorithm ([31], [61]) was successfully ap-
plied on acquired data. Indeed, an experimental study had to be set up to get a
precise dynamical evaluation of the disease area on the leaves. The posterior dis-
495 tributions obtained for the three parameters allowed us to update our knowledge
on the parameters of the Macho por Hembra plantain variety. The statistical
model used to define the likelihood contains a residual error variance, which in
this approach is considered as an unknown parameter and was estimated. The
assumption about the independent residual errors could be improved to bet-
500 ter describe all the possible sources of model simulation uncertainty. We could
try to define a more realistic statistical model that takes into account residual
variance between leaf compartments and residual variance within leaf compart-
ments (considered as time variability). The actual residual error variance could
be defined as the sum of a term describing the variability between leaf compart-
505 ments and a term describing the variability between observation dates. The
SiBaToka model was validated at plant scale, during the two first plant cycles
(630 days) on a susceptible plantain variety. Simulations were consistent with
experimental data. Indeed, the calculated root mean square error was 9.59%.
The model built at leaf compartment scale and the validation at plant cycle
510 scale show the robustness of the predictions.

However, the necrotic surface per plant simulated by the model was fre-
quently lower than the observed necrotic surface (Figure 3). These differences
could be explained by a varietal effect and/or some climatic differences inducing
differences in inoculum pressure between Dominican Republic where the data
515 for calibration were acquired on the variety Macho por Hembra and Cameroon
where the data for validation were obtained on the variety French Sombre. In-
deed, the average rainfalls are 1850 mm and 2400 mm respectively for Dominican
Republic and Cameroon. A decreasing disease severity was observed between
481 to 571 days on the data whereas we noted a stable severity during this pe-

520 riod on simulations. This BLSD decrease can be explained by a natural decrease
of inoculum pressure due to the occurrence of dry season in Cameroon whereas
the model supposes a stable inoculum in the air within time and cycles.

Simulations with inferred parameters allowed to improve our knowledge on
the BLSD epidemiology at plant scale and the effect of banana growth on disease
525 level. The simulations on three different varieties showed the same range of
susceptibility to BLSD as [25] measured under field conditions, and evaluated
the variety Pisang Berlin more resistant than Pisang Madu. The effect of lesion
growth on severity showed that the lowest severity level was obtained on plant
B with the lowest lesion growth rate. However the severity simulated on the
530 variety Pisang madu-like is higher than expected according to our expertise on
this variety. Parameter values used in the plant sub-model and some values of
the other epidemiological parameters in the model could explained this result.

Even if the model of architecture of the banana is simple, it allows to cal-
culate the approximate distances between leaf compartments and so to model
535 spore dispersal at plant scale. The model allowed to observe a slight effect of
the distance of dispersal of asexual spores on the asexual sporulating surface
(Figure 5) but not on the necrotic surface. The effect of spore dispersal on the
density of news infections was also detected on foliar diseases of wheat ((Maae-
nen, EJPP 2003). A global analysis of sensitivity of the model output could
540 confirm this result.

At least, simulations allowed to show that a high leaf emission rate can
decrease severely the disease severity up to 80% (Figure 6). [18] already observed
such effect in Nigeria in experimentations. Nevertheless, the SiBaToka model
does not include any interaction between disease and host growth whereas some
545 authors mentioned several impact of BLSD on banana growth like a delay of
one month for flowering date on plantains infected by BLSD [44].

The model will be a tool to better understand the BLSD dynamics at plant
scale, particularly concerning autoinfection. Indeed, it is very difficult to quan-
tify the secondary inoculum at plant scale. On BLSD, few studies quantified
550 the dynamics of spore concentrations at different heights with spore traps and

these studies have not showed any correlation between spore densities in the air and disease severity at plant scale ([29] [51])

Some generic model of host-pathogen including host growth have been recently developed [13] but they were not adapted to the specificity of bananas and BLSD. As, the model SiBaToka includes a specific pathogen epidemic cycle, with both asexual and sexual reproduction, and a more generic plant sub-model defined by leaf emission. The model was defined in optimal climatic condition because of the presence of the pathosystem only in tropical and sub-tropical climates: it doesn't take account of the changes in temperature. It would be interesting to add climatic variables in the model that could reduce or accelerate the epidemic because it is established that *M.Fijiensis* development is climate dependent [15]. Another way to take account of the variability in disease development will be to improve the dependence between input parameters, as described in [44], using temporal variability in parameters definition [19]. In the context of BLSD invasion, another extend of the model will be to integrate the competition with the pathogen *Mycosphaerella Musicola* responsible of the Sigatoka disease. Indeed, Sigatoka disease was present in any banana producing area, and after BLSD arrived, it has always substituted Sigatoka disease after several years.

The model developed here will be used as a tool to understand the BLSD epidemic and will be very helpful to design and evaluate efficient control practices.

As a perspective, it will be interesting to increase the spatial scale of the model, at plot scale for example, by modelling the development of the pathosystem at plant scale and including the interactions between plants.

5. Acknowledgments

We would like to thank Virginie Ravigné, Jean Carlier, Marc Dorel and David Makowski for their scientific discussion on the model, Teresa Martinez and Salomon Emmanuel from Instituto Dominicano de Investigaciones Agropecuar-

580 ias y Forestales for the realization of the experimentation in Dominican Republic
 and Aurore Cavalier and Yolande Chilin-Charles for technical assistance. We
 thank Thomas Merle from Fédération Régionale d’Eradication et de Destruction
 des Organismes Nuisibles for his assistance for the experimentation in Guade-
 loupe. We thank also CARBAP (Cameroon) and CIRAD (Montpellier) for the
 585 realization of the past experimentations. We are particularly grateful to Raphael
 Pasquier of University of Antilles-Guyane (C3I) for his computing assistance
 for the numerical experimentations and we thank Samuel Ralambondrainy and
 Jean-Christophe Guenot for their assistance in the image analysis. Clara Landry
 conducted her PhD which was supported by a doctoral fellowship from CIRAD.
 590 The fellowship and the work were financially supported by FEDER Guade-
 loupe through a research project (the convention 1/1.4-30305project/Gestion
 des risques en santé animale et végétale) and a cooperation project through the
 Interreg IV Caraibes fundings (convention num. 31409/CABARé).

References

- 595 [1] Abadie, C., Chilin-Charles, Y., Huat, J., Salmon, F., Pignolet, J., Carlier,
 J., Lescot, T., Côte, F., & Jenny, C. (2009). New approaches to select
 durable resistant varieties of banana against mycosphaerella leaf spot dis-
 eases. *Acta Horticulturae*, 828, 171–178.
- 600 [2] Abadie, C., Zapater, M. F., Pignolet, L., Carlier, J., & Mourichon, X.
 (2008). Artificial inoculation on plants and banana leaf pieces with my-
 cosphaerella spp., responsible for sigatoka leaf spot diseases. *Fruits*, 63,
 319–323.
- [3] Amil, A., Heaney, S., Stanger, C., & Shaw, M. W. (2007). Dynamics of qoi
 sensitivity in mycosphaerella fijiensis in costa rica during 2000 to 2003g.
 605 *Phytopathology*, .
- [4] Arneson, P. (1989). A mechanistic computer simulation model of sigatoka.
 In *Sigatoka leaf spot diseases of bananas*.

- [5] Audsley, E., Milne, A., & Paveley, N. (2005). A foliar disease model for use in wheat disease management decision support systems. *Ann. Appl. Biol*, 147, 161–172.
- [6] Bakry, F., Careel, F., Jenny, C., & Horry, J. (2009). Breeding plant tree crop: tropical species. chapter Genetic improvement of banana. (pp. 3–50). Springer[US].
- [7] van den Berg, F., Gaucel, S., Lannou, C., Gilligan, C. A., & van den Bosch, F. (2013). High levels of auto-infection in plant pathogens favour short latent periods: a theoretical approach. *Evolutionary Ecology*, 27, 409–428.
- [8] Brisson, N., Ozier-Lafontaine, H., & Dorel, M. (1998). Effects of soil management and water regime on banana growth between planting and flowering simulation using the stics model. *Acta Horticulturae*, 490.
- [9] Burt, P. J. A., Rosenberg, L. J., Rutter, J., Ramirez, F., & Gonzales O, H. (1999). Forecasting the airborne spread of mycosphaerella fijiensis a cause of black sigatoka disease on banana: estimations of numbers of perithecia and ascospores. *Annals of Applied Biology*, 135, 369–377.
- [10] Burt, P. J. A., Rutter, J., & Ramirez, F. (1998). Airborne spore loads and mesoscale dispersal of the fungal pathogens causing sigatoka diseases in banana and plantain. *Aerobiologia*, 14, 209–214.
- [11] Calonnec, A., Cartolaro, P., Naulin, J.-M., Bailey, D., & Langlais, M. (2008). A host-pathogen simulation model: powdery mildew of grapevine. *Plant Pathology*, 57, 493–508.
- [12] Carlier, J., Mourichon, X., & Jones, D. (2000). Diseases of banana, abaca and enset. chapter Black leaf streak: the causal agent. (pp. 48–56). CABI Publishing [New York].

- [13] Casadebaig, P., Quesnel, G., Langlais, M., & Faivre, R. (2012). A generic
635 model to simulate air-borne diseases as a function of crop architecture.
PLoS ONE, 7, 1–11.
- [14] Caubel, J., Launay, M., Lannou, C., & Brisson, N. (2012). Generic response
functions to simulate climate-based processes in models for the development
of airborne fungal crop pathogens. *Ecological Modelling*, 242, 92–104.
- 640 [15] Churchill, A. (2010). *Mycosphaerella fijiensis*, the black leaf streak
pathogen of banana: progress towards understanding pathogen biology and
detection, disease development, and the challenges of control. *Molecular
plant pathology*, (pp. 1–22).
- [16] Cohan, J., Abadie, C., Tomekpe, K., & J., T. T. (2003). Evaluation des
645 performances agronomiques et de rsistance la maladie des raies noires
de l’hybride de plantain crbp 39 du carbap. rsultats du premier cycle de
culture. *Infomusa*, 1.
- [17] Costes, E., Lauri, P. E., Simon, S., & Andrieu, B. (2013). Plant architec-
ture, its diversity and manipulation in agronomic conditions, in relation
650 with pest and pathogen attacks. *Eur J Plant Pathol*, .
- [18] Craenen, K., & Ortiz, R. (1998). Influence of black sigatoka disease on
the growth and yield of diploid and tetraploid hybrid plantains. *Crop
Protection*, 17, 13–18.
- [19] Cunniffe, N. J., & Gilligan, C. A. (2010). Invasion, persistence and
655 control in epidemic models for plant pathogens: the effect of host
demography. *Journal of The Royal Society Interface*, 7, 439–451.
arXiv:<http://rsif.royalsocietypublishing.org/content/7/44/439.full.pdf+html>.
- [20] De Wolf, E. D., & Isard, S. A. (2007). Disease cycle approach to plant
disease prediction. *Ann. Rev. Phytopathology*, 45, 203–220.

- 660 [21] Donzelli, B., & Churchill, A. (2007). A quantitative assay using mycelial fragments to assess virulence of *mycosphaerella fijiensis*. *Phytopathology*, 97-8, 916–929.
- [22] Ferrandino, F. J. (2008). Effect of crop growth and canopy filtration on the dynamics of plant disease epidemics spread by aerially dispersed spores.
665 *Phytopathology*, 98, 492–503.
- [23] Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., & Gurr, S. J. (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature*, 484, 186–194.
- [24] Fouré, E. (1982). Les cercosporioses du bananier et leurs traitements. comportement des variétés. etude de la sensibilité variétale des bananiers et
670 plantain à *mycosphaerella fijiensis* morelet au Gabon (maladie des raies noires). *Fruits*, 37, 749–766.
- [25] Fouré, E. (1990). Etude de la sensibilité variétale des bananiers et plantain à *mycosphaerella fijiensis* morelet au cameroun. caractérisation de la
675 résistance au champ de bananiers appartenant à divers groupes génétiques. *Fruits*, 45, 339–345.
- [26] Fouré, E., Grisoni, M., & Zurfluh, R. (1984). Les cercosporioses du bananier et leurs traitements. comportement des variétés. *Fruits*, 39, 365–378.
- [27] Fouré, E., & Moreau, A. (1992). Contribution à l'étude épidémiologique de la cercosporiose noire dans la zone bananière du moungo au cameroun de
680 1987 à 1989. *Fruits*, 47, 3–16.
- [28] Garin, G., Fournier, C., Andrieu, B., Houls, V., Robert, C., & Pradal, C. (2014). A modelling framework to simulate foliar fungal epidemics using functionalstructural plant models. *Annals of Botany*, 114, 795–812.
- 685 [29] Gauhl, F. (1994). *Epidemiology and Ecology of Black Sigatoka (Mycosphaerella fijiensis Morelet) on Plantain and Banana (Musa spp.) in Costa Rica, Central America..* inibap.

- [30] Gelman, A., & Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–511.
- 690 [31] Gouache, D., Bensadoun, A., Brun, F., Pagé, C., Makowski, D., & Wallach, D. (2012). Modelling climate change impact on septoria tritici blotch (stb) in France: Accounting for climate model and disease model uncertainty. *Agricultural and Forest Meteorology*, .
- 695 [32] Irish, B., Goenaga, R., Rios, C., Chavarria-Carvajal, J., & Ploetz, R. (2013). Evaluation of banana hybrids for tolerance to black leaf streak (mycosphaerella fijiensis morelet) in Puerto Rico. *Crop Protection*, 54, 229–238.
- [33] Jeger, M. (2004). Analysis of disease progress as a basis for evaluating disease management practice. *Ann. Rev. Phytopathol.*, 42, 61–82.
- 700 [34] Jones, D. (2000). *Diseases of banana, abaca and enset*. CABI publishing.
- [35] Lannou, C. (2012). Variation and selection of quantitative traits in plant pathogens. *Annual Review of Phytopathology*, 50, 319–338. PMID: 22702351.
- 705 [36] Lannou, C., Soubeyrand, S., Frezal, L., & Chaduf, J. (2008). Autoinfection in wheat leaf rust epidemics. *New Phytologist*, 177, 1001–1011.
- [37] Lassoudière, A. (2006). *Le bananier et sa culture*. Quae Editions.
- [38] Lassoudière, A. (2012). *Le bananier, Un siècle d’innovations techniques..* Editions Quae.
- 710 [39] Lo Iacono, G., van den Bosch, F., & Gilligan, C. A. (2013). Durable resistance to crop pathogens: An epidemiological framework to predict risk under uncertainty. *PLoS Comput Biol*, 9.
- [40] Lô Pelzer, E., Bousset, L., Jeuffroy, M., Salam, M., Pinochet, X., Boillot, M., & Aubertot, J. (2010). Sippom-wosr: A simulator for integrated

- pathogen population management of phoma stem canker on winter oilseed
715 rape: I. description of the model. *Field Crops Research*, 118, 73–81.
- [41] Marin, D. H., Romero, R. A., Guzman, M., & Sutton, T. B. (2003). Black
sigatoka: An increasing threat to banana cultivation. *Plant Disease*, 87,
208–222.
- [42] Meredith, D., & Lawrence, J. (1969). Black leaf streak disease of bananas
720 (mycosphaerella fijiensis): Symptoms of disease in hawaii, and notes on the
conidial state of the causal fungus. *Transactions of the British Mycological
Society*, 52, 459 – IN16.
- [43] Milne, A., Paveley, N., Audsley, E., & Livermore, P. (2003). A wheat
canopy model for use in disease management decision support systems.
725 *Ann. appl. Biol.*, 143, 265–274.
- [44] Mobambo, K., Gauhl, F., Vuylsteke, D., Ortiz, R., Pasberg-Gauhl, C., &
Swennen, R. (1993). Yield loss in plant from black sigatoka leaf spot and
field performance of resistant hybrids. *Field Crops Research*, 35, 35–42.
- [45] Nakagawa, Y., Yokozawa, M., & Hara, T. (2015). Competition among
730 plants can lead to an increase in aggregation of smaller plants around larger
ones. *Ecological Modelling*, 301, 41 – 53.
- [46] Papaïx, J. (2011). *Structure du paysage agricole et risque épidémique, une
approche démo-génétique..* Ph.D. thesis AgroParisTech.
- [47] Rieux, A., Soubeyrand, S., Bonnot, F., Klein, E. K., Ngando, J. E., Mehl,
735 A., Ravigne, V., Carlier, J., & de Lapeyre de Bellaire, L. (2014). Long-
distance wind-dispersal of spores in a fungal plant pathogen: estimation of
anisotropic dispersal kernels from an extensive field experiment. *Plos One*,
.
- [48] Robert, C., & Casella, G. (2010). *Monte Carlo Statistical Methods*. second
740 ed. Springer.

- [49] Robert, C., Fournier, C., Andrieu, B., & Ney, B. (2008). Coupling a 3d virtual wheat (*triticum aestivum*) plant model with a septoria tritici epidemic model (septo3d): a new approach to investigate plant-pathogen interactions linked to canopy architecture. *Funct. Plant Biol.*, *35*, 997–1013.
- 745 [50] Robinson, J. (2003). *Bananas and Plantains*. CAB INTERNATIONAL.
- [51] Rocha, H., Pozza, E., Uchoa, C., Cordeiro, J., De Souza, P., Sussel, A., & De Rezende, C. (2012). Temporal progress of yellow sigatoka and aerobiology of mycosphaerella musicola spores. *Journal of phytopathology*, *160*.
- [52] Rossi, V., Caffi, T., Giosuè, S., & Bugiani, R. (2008). A mechanistic model
750 simulating primary infections of downy mildew in grapevine. *Ecological Modelling*, *212*, 480–491.
- [53] Rossi, V., & Giosuè, S. (2003). A dynamic simulation model for powdery mildew epidemics on winter wheat. *EPPO Bulletin*, *33*, 389–396.
- [54] Rossi, V., Giosue, S., & Caffi, T. (2009). Modelling the dynamics of in-
755 fections cause by sexual and asexual spores during plasmopara viticola epidemics. *Journal of Plant Pathology*, *91*, 615–627.
- [55] Rossi, V., Racca, P., Giosue, S., Pancaldi, D., & Alberti, I. (1997). A simulation model for the development of brown rust epidemics in winter wheat. *European Journal of Plant Pathology*, *103*, 453–465.
- 760 [56] Sackett, K., & Mundt, C. (2005). The effects of dispersal gradient and pathogen life cycle components on epidemics velocity in computer simulations. *The american phytopathological society*, *95*, 992–1000.
- [57] Saltelli, A., Tarantola, S., & K.P.S., C. (1999). A quantitative model independent method for global sensitivity analysis of model output. *Techno-
765 metrics*, *41*, 39–56.
- [58] Stover, R. (1980). Sigatoka leaf spots of bananas and plantains. *Plant disease*, *64*.

- [59] Stover, R., & Simmonds, N. (1987). *Bananas*. Longman Scientific & Technical, third edition.
- 770 [60] Tixier, P., Malezieux, E., & Dorel, M. (2004). Simba-pop: a cohort population model for long-term simulation of banana crop harvest. *Ecological Modelling*, 180, 407–417.
- [61] Wagner, F., Hérault, B., Stahl, C., Bonal, D., & Rossi, V. (2011). Modeling water availability for trees in tropical forests. *Agricultural and Forest*
775 *Meteorology*, 151, 1202–1213.
- [62] Wallach, D. (2006). Evaluating crop models. chapter 2. (pp. 11–54). Elsevier.
- [63] Willocquet, L., & Savary, S. (2004). An epidemiological simulation model with three scales of spatial hierarchy. *Phytopathology*, 94, 883–891.
- 780 [64] Xu (1999). Modelling and forecasting epidemics of apple powdery mildew (*podosphaera leucotricha*). *Plant Pathology*, 48, 462–471.

6. Appendix A

Consider a foliar surface where lesions grow and can coalesce. We suppose that (i) lesions are elliptic, with the same ratio (minor axis)/(major axis) and with parallel major axes, (ii) the centers of the lesions are randomly distributed and spatially independent, and (iii) lesions fall into k size classes such that n_i lesions of surface area s_i fall in class i . We make the approximation that the sporulating area is the sum of the areas of overlapping compatible lesions. Let a_i and b_i be the lengths of the semi-major and semi-minor axes of an ellipse in class i . Consider two classes i and j . Any ellipse E_j in class j intersects a particular ellipse E in class i if the center of E_j falls inside the ellipse E' that has the same center as E and whose semi-major and semi-minor axis lengths equal respectively $(a_i + a_j)$ and $(b_i + b_j)$. Neglecting the edge effect of the foliar surface, the probability p_{ij} that the center of E_j falls inside E' equals approximately the ratio of the area of E' over the area A of the foliar surface, i.e. $p_{ij} = \pi(a_i + a_j)(b_i + b_j)/A$. The probability that no ellipse in class j intersects E is therefore $(1 - p_{ij})^{n_j}$ if $i \neq j$ and $(1 - p_{ii})^{n_i - 1}$ if $i = j$ (because there are $n_i - 1$ ellipses excluding E in class i). The probability for an ellipse in class i to be intersected by at least another ellipse is therefore $q_i = 1 - (1 - \pi(2a_i)(2b_i)/A)^{n_i - 1} \prod_{j, j \neq i} (1 - \pi(a_i + a_j)(b_i + b_j)/A)^{n_j}$. Given $s_i = \pi a_i b_i$ and $b_i/a_i = b_j/a_j$, the equation becomes $q_i = 1 - (1 - 4s_i/A)^{n_i - 1} \prod_{j, j \neq i} (1 - (\sqrt{s_i} + \sqrt{s_j})^2/A)^{n_j}$.

7. Appendix B

The Metropolis-Hastings algorithm within Gibbs algorithm used to approach
 805 the posterior distribution of parameters $\theta = \{\theta_1, \theta_2, \theta_3\}$ and residual variance σ^2 , works as follow.

Step 0: parameters initialisation.

- update $\theta^0 = \{\theta_1^0, \theta_2^0, \theta_3^0\}$
- update $\sigma^{2(0)} \sim \Gamma^{-1}(\alpha, \beta)$ with
 810 $\alpha = (3 \times 21 \times 3 \times 7)/2 + 2$
 $\beta = \frac{\sum_{p=1}^3 \sum_{t=1}^{21} \sum_{\ell=1}^3 \sum_{c=1}^7 [y_{\ell,c}^p(t) - Y_{\ell,c}^p(t)]^2}{2}$

Step 1: Metropolis-Hastings step. Example for the iteration n for the first parameter parameter.

- proposition $\tilde{\theta}_1 \sim \text{prop}_{\theta_1}(\theta_1^{n-1})$ we have

$$\tilde{\theta} = \{\tilde{\theta}_1, \theta_2^{n-1}, \theta_3^{n-1}\}$$

- acceptance or rejection of the new candidate $\tilde{\theta}_1$ by computing

$$\rho = \frac{\pi(Y|\tilde{\theta}, \sigma^{2(n-1)}) \pi(\tilde{\theta}_1)}{\pi(Y|\theta^{n-1}, \sigma^{2(n-1)}) \pi(\theta_1^{n-1})} \times \frac{\text{prop}_{\theta}(\theta_1^{n-1}|\tilde{\theta}_1)}{\text{prop}_{\theta}(\tilde{\theta}_1|\theta_1^{n-1})}$$

$$\theta_1^n = \begin{cases} \tilde{\theta}_1 & \text{si } u \sim U([0, 1]) \leq \rho \\ \theta_1^{n-1} & \text{else} \end{cases}$$

Repeat this *Step 1* for the second and the third parameter.

815 *Step 2: Gibbs step.*

- update $\sigma^{2(n)} \sim \Gamma^{-1}(\alpha, \beta)$ with:
 $\alpha = (3 \times 21 \times 3 \times 7)/2 + 2$
 $\beta = \frac{\sum_{\ell=1}^3 \sum_{t=1}^{21} \sum_{\ell=1}^3 \sum_{c=1}^7 [y_{\ell,c}^p(t) - Y_{\ell,c}^p(t)]^2}{2}$

Repeat *Step 1* and *Step 2* until convergence.

We give proposal distribution and boundaries of the distribution:

$$prop_{lgr} = \mathcal{N}(lgr^{n-1}, 0.004); \text{ truncated on } [0, 1]$$

$$prop_{ei} = \mathcal{N}(ei^{n-1}, 0.001); \text{ truncated on } [0, 0.1]$$

$$prop_{di} = \mathcal{N}(di^{n-1}, 0.8); \text{ truncated on } [0, 50]$$

Calculus explained for the posterior distribution of σ^2 in the Gibbs step. Let

$\tau = 1/\sigma^2$ and $N = 3 \times 21 \times 3 \times 7$ be the number of observations, we have

$$\pi(\tau|Y, \theta) \propto \tau^{((N/2)+1)} \times \exp(-\tau \sum_{p=1}^3 \sum_{t=1}^{21} \sum_{\ell=1}^3 \sum_{c=1}^7 (y_{\ell,c}^p(t) - Y_{\ell,c}^p(t))^2/2)$$

We recognize a Gamma distribution:

$$\pi(\tau|Y, \theta) \sim \Gamma\left(\frac{N}{2} + 2, \quad 2 / \sum_{p=1}^3 \sum_{t=1}^{21} \sum_{\ell=1}^3 \sum_{c=1}^7 (y_{\ell,c}^p(t) - Y_{\ell,c}^p(t))^2\right)$$

and

$$\pi(\sigma^2|Y, \theta) \sim \Gamma^{-1}\left(\frac{N}{2} + 2, \quad \sum_{p=1}^3 \sum_{t=1}^{21} \sum_{\ell=1}^3 \sum_{c=1}^7 (y_{\ell,c}^p(t) - Y_{\ell,c}^p(t))^2/2\right)$$

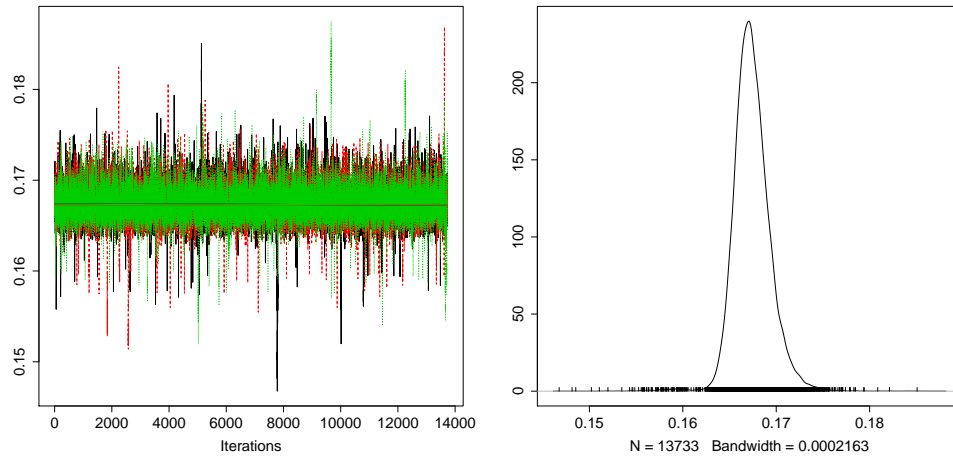


Figure 7: Trace of lgr parameter markov chain for the three algorithm runs (color black, red, green) and posterior density of the parameter. A burn in of 1000 iterations and a thinning of 5 was used on the 70000 iterations of each run.

2.2 Simulation et évaluation du modèle

Dans cette partie les valeurs des paramètres sont fixées aux valeurs moyennes décrites dans [Landry et al. 2015, in prep.] présenté précédemment. Les paramètres du module plante et du module épidémiologique sont récapitulés dans les tableaux 2.1 et 2.2.

Paramètre	Description	Valeur moyenne	Unité	Référence
Δ_{fh}	temps entre la floraison et la récolte	90	jours	Irish et al. [2013]
ϕ_L^v	angle entre deux feuilles consécutives (vertical)	10	degrés	Stover and Simmonds [1987]
L_L^{max}	longueur de feuille maximale	2016	mm	Stover and Simmonds [1987] [Landry et al., in prep.]
L_L^{min}	longueur de feuille minimale	486	mm	[Landry et al., in prep.]
W_L^{max}	largeur de feuille maximale	846	mm	Stover and Simmonds [1987]
W_L^{min*}	largeur de feuille minimale	253	mm	[Landry et al., in prep.]
ϕ_L^H	angle phyllotaxique	156	degrés	Lassoudière [2012]
l_p^0	rapport longueur pétiole/feuille à la plantation	0.162	-	[Landry et al., in prep.]
l_p^1	rapport longueur pétiole/feuille à l'apparition du rejet	0.238	-	[Landry et al., in prep.]
H_L^0	hauteur du point d'insertion de la feuille la plus basse à la plantation	360	mm	[Landry et al., in prep.]
N_L^{flower}	nombre de feuilles émises avant la floraison	40	-	Robinson [2003]
H_L^1	hauteur du point d'intersection de la feuille la plus basse à l'apparition du rejet	449	mm	[Landry et al., in prep.]
N_L^0	nombre de feuilles émises à l'apparition du rejet sélectionné	18	-	Lassoudière [2006]
l_{node}	longueur des entrenœuds	87	mm	Robinson [2003]
ler^{init}	rythme d'émission foliaire initial	1.32	feuille /semaine	Stover and Simmonds [1987]
ler^{final}	rythme d'émission foliaire final	1	feuille /semaine	Stover and Simmonds [1987]
$N_L^{harvest}$	nombre de feuilles sur la plante à la récolte	5	-	Irish et al. [2013]
N_L^{max}	nombre de feuilles maximal sur la plante	12	-	Irish et al. [2013]
$N_L^{0,init}$	nombre de feuilles sur la plante à la plantation	3	-	[Landry et al., in prep.]
$N_L^{1,init}$	nombre de feuilles du rejet à la récolte du pied mère	10	-	Lassoudière [2012]

TABLEAU 2.1 – Paramètres du module plante.

Paramètre	Description	Valeur moyennes	Unité	Référence
ei^*	efficacité d'infection sur une feuille nouvellement émise (nombre de lésion crée par une spore)	0.0167	-	[Landry et al., in prep.]
ip^*	durée d'incubation : temps entre l'infection et l'apparition de la lésion	31	jours	Fouré [1982]
lgr^*	taux de croissance des lésions	0.1674	mm ² /jour	[Landry et al., in prep.]
ei^{age}	effet de l'âge de la feuille sur l'efficacité d'infection	0.08	-	-
a^{min}	taille minimale d'une lésion capable d'émettre des spores asexuées	2	mm ²	Fouré [1982]
sc_a	capacité de sporulation asexuée	7	nombre /mm ²	Fouré [1982]
sc_s	capacité de sporulation sexuée	182	nombre /mm ²	Fouré et al. [1984] Burt et al. [1999]
Δ_a	durée de sporulation asexuée	8	jours	[Landry et al., in prep.]
Δ_s	durée de sporulation sexuée	1	jours	Jones [2000] Gauhl [1994]
Δ_{a-s}	durée entre la sporulation asexuée et sexuée	10	jours	Abadie, pers.data
β_a	paramètre de décroissance de la fonction de dispersion des spores asexuées	0.001	-	Rieux et al. [2013]
β_s	paramètre de décroissance de la fonction de dispersion des spores sexuées	0.00001	-	Rieux et al. [2013]

TABEAU 2.2 – Paramètres du module épidémiologique. Les paramètres notés * correspondent aux valeurs moyennes des densités a posteriori des paramètres inférés.

2.2.1 Description de la plante

La surface foliaire. Elle est définie par des longueurs et largeurs de feuilles. Au cours du temps les feuilles sont émises avec des dimensions de plus en plus importantes mais qui ne changent pas au cours de la vie de la feuille. L'augmentation avec le temps de la largeur des feuilles émises peut être modélisée par une fonction croissante avec plateau. La surface foliaire totale de la plante croît au cours du cycle de la plante et est modélisée par la relation décrite par [Lassoudière, 2006] (Figure 2.1 et 2.2). La surface d'une feuille varie de 0,1 m² à la plantation à 1,7 m² à la floraison.

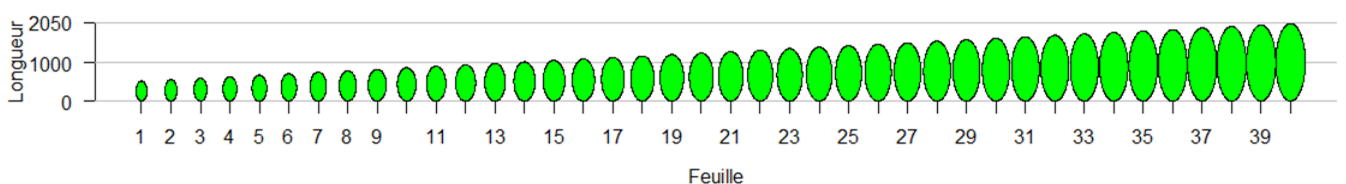


FIGURE 2.1 – Représentation schématique des dimensions des 40 feuilles émises au cours d'un cycle de culture.

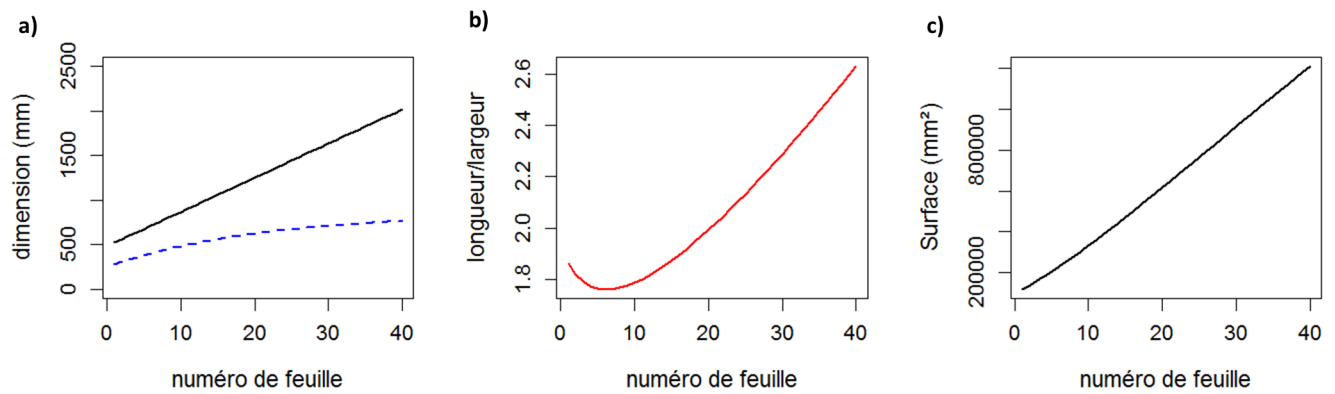


FIGURE 2.2 – Évolution dans le temps des dimensions des feuilles émises en fonction du numéro la de feuille émise. a) longueurs (trait plein) et largeurs (pointillé), b) ratio longueur/largeur et c) surface totale.

Dans le modèle, chaque feuille est divisée en sept compartiments correspondant aux étapes du déroulement de la première feuille ou cigare (Figure 2.3 a). Nous avons réalisé une modélisation simplifiée de l'architecture du bananier qui prend en compte la hauteur du point d'insertion des feuilles sur le pseudo-tronc et leur inclinaison en fonction du temps (Figure 2.3b). Il est donc aisé de calculer les coordonnées dans l'espace des centres des compartiments et les distances qui les séparent.

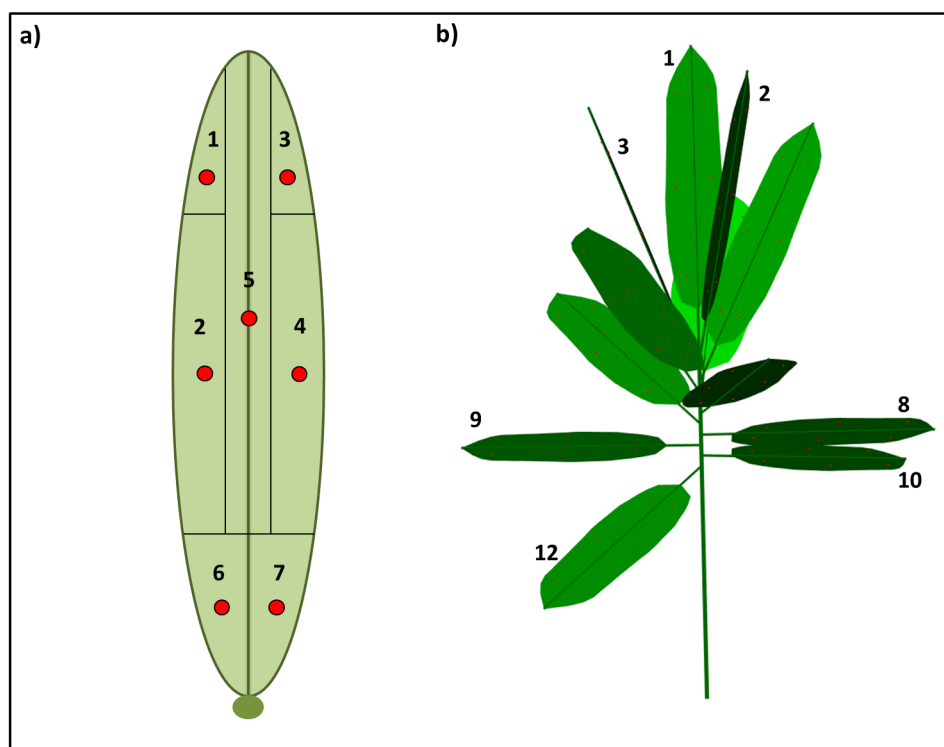


FIGURE 2.3 – Représentations schématiques : a) Division des feuilles en sept compartiments, les points rouges représentent les centres des compartiments. b) représentation en 3 dimensions d'un bananier au 170 ième jour de croissance avec 12 feuilles numérotées de la plus jeune (feuille 1) à la plus ancienne (feuille 12)

Les cycles de cultures. Ils sont tous décrits de manière identique, à l'exception du premier qui a ses caractéristiques propres car il débute à la plantation (Figure 2.4). La modélisation de la croissance des rejets tient compte de l'inhibition de leur croissance par le bananier du cycle précédent. En début de deuxième cycle, la surface foliaire augmente moins rapidement qu'en premier cycle et à la récolte du pied mère, les surfaces foliaires deviennent identiques d'un cycle à l'autre.

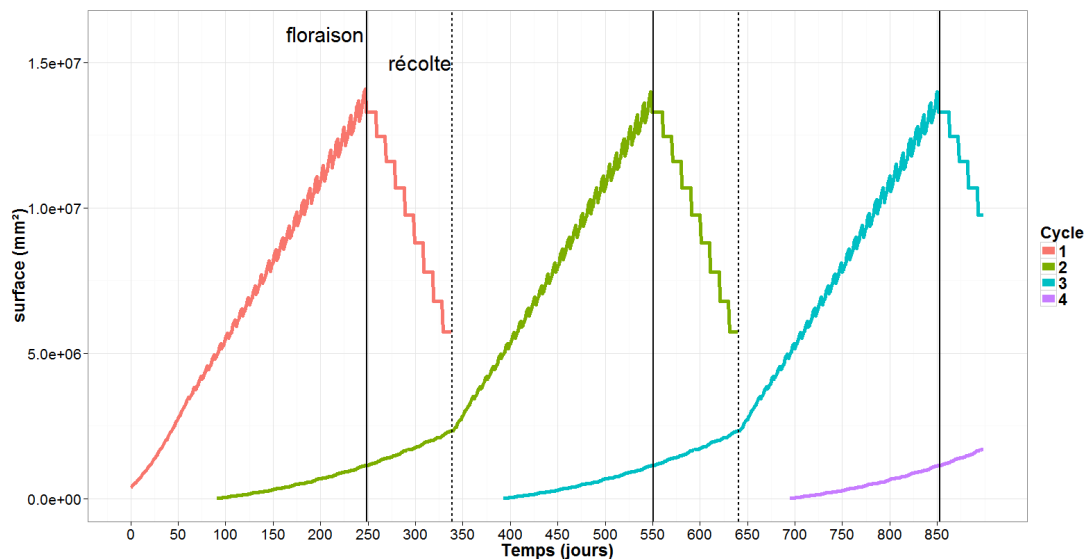


FIGURE 2.4 – Dynamique de la surface foliaire au cours de trois cycles de culture. Les barres verticales correspondent respectivement aux dates de floraison et aux dates de récolte.

2.2.2 Densité d'inoculum primaire

Afin d'initier une épidémie, et sachant les paramètres relatifs à une variété de type plantain (Tableaux 2.1 et 2.2), nous avons estimé par simulations que la quantité minimale d'inoculum primaire nécessaire à l'initiation de l'épidémie est comprise entre 0.0001 spores/mm²/jour, densité avec laquelle l'épidémie ne s'initie jamais, et 0.0002 spores/mm²/jour, densité pour laquelle une lésion en incubation apparaît à l'ouverture de la 30^{ème} feuille émise du premier cycle. Compte tenu de la durée d'incubation du pathogène, l'apparition du premier symptôme de la maladie ne se fait qu'après la floraison.

Pour une autre variété sensible décrite avec des paramètres épidémiologiques proche de ceux d'une variété de type Cavendish (banane dessert largement utilisée dans la production pour l'exportation) : $sc_a = 92$, $sc_s = 182$, $\Delta_a = 12$, $di = 18$, $lgr = 0.12$ et $ei = 0.009$. La quantité minimale de spores présente dans l'inoculum journalier nécessaire à l'initiation de l'épidémie est comprise entre 0.0002 spores/mm²/jour et 0.0003 spores/mm²/jour. La première infection (lésion en incubation) apparaît sur la 34^{ème} feuille émise après la plantation.

L'épidémie initiée, le modèle simule sur plusieurs cycles de culture des pourcentages de surface atteinte par la cercosporiose noire à différents stades du cycle épidémique. La Figure 2.5 reproduit la dynamique de quatre sorties d'intérêt, sur trois cycles de culture. Le premier cycle dure

de la plantation à la récolte, le second et le troisième cycle sont plus long : le rejet se développe lentement jusqu'à la récolte du plant précédent puis son rythme d'émission foliaire augmente entraînant une diminution du pourcentage de maladie totale et de nécrose jusqu'à l'apparition du rejet suivant.

Il est intéressant de noter que, dans le cas proposé en simulation (Figure 2.5), plus de la moitié de la surface malade correspond à la surface nécrosée et que la surface conidienne reste inférieure à 5 %.

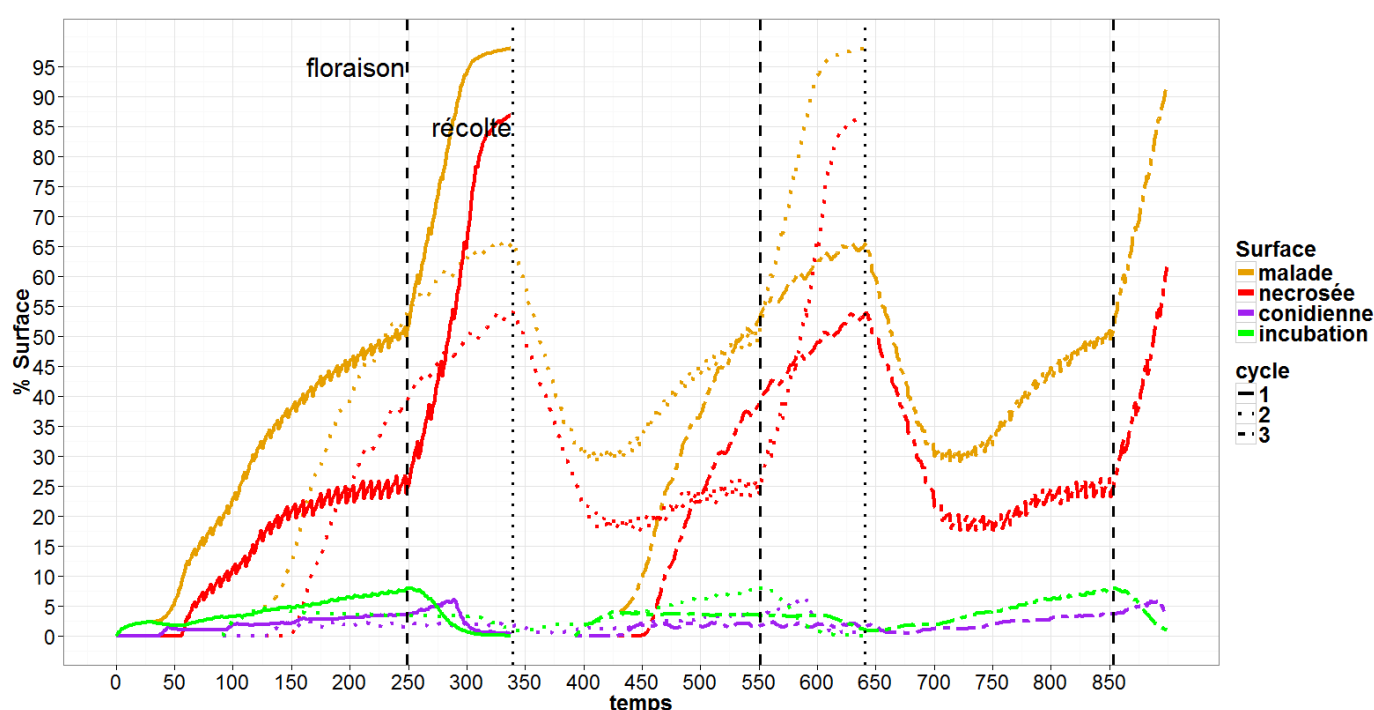


FIGURE 2.5 – Dynamique de la surface malade au cours de trois cycles de culture initiée avec 0.2 spores/mm²/jour. Les barres verticales correspondent respectivement aux dates de floraison, et aux dates de récolte.

Le modèle permet aussi de simuler sur plusieurs cycles de culture les nombres de lésions créées suivant leurs états épidémiologiques : en incubation, conidiennes (sporulantes en spores asexuées), nécrosées (sporulantes en spores sexuées) (Figure 2.6). Ces simulations permettent de comparer les quantités des différents types de lésions au cours du cycle phénologique de la plante.

La dynamique des différents types de lésions (Figure 2.6) montre que le nombre de lésions total est maximal à la floraison et atteint plus de 2 millions de lésions/plant dont plus de la moitié sont des lésions invisibles, des lésions en incubation. Parmi les autres lésions provoquant des symptômes sur feuilles (lésions conidiennes et nécrosées), les lésions nécrosées représentent plus de 50% de l'effectif et sont très majoritaires après la floraison.

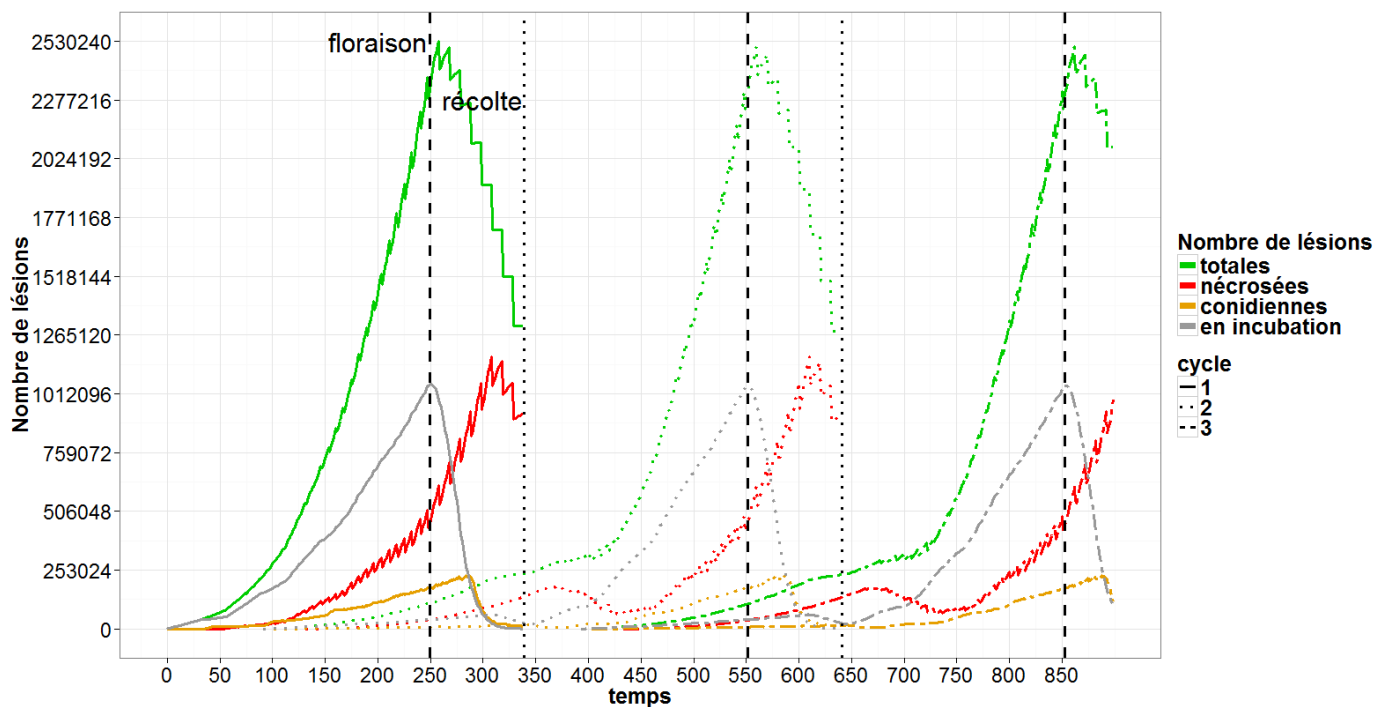


FIGURE 2.6 – Dynamique du nombre de lésions au cours de trois cycles de culture initiée avec 0.2 spores/mm²/jour. Les barres verticales correspondent aux dates de floraison (tirets) et aux dates de récolte (pointillés).

De manière plus précise, il nous est possible de connaître la dynamique de la maladie sur chacune des feuilles (Figure 2.7).

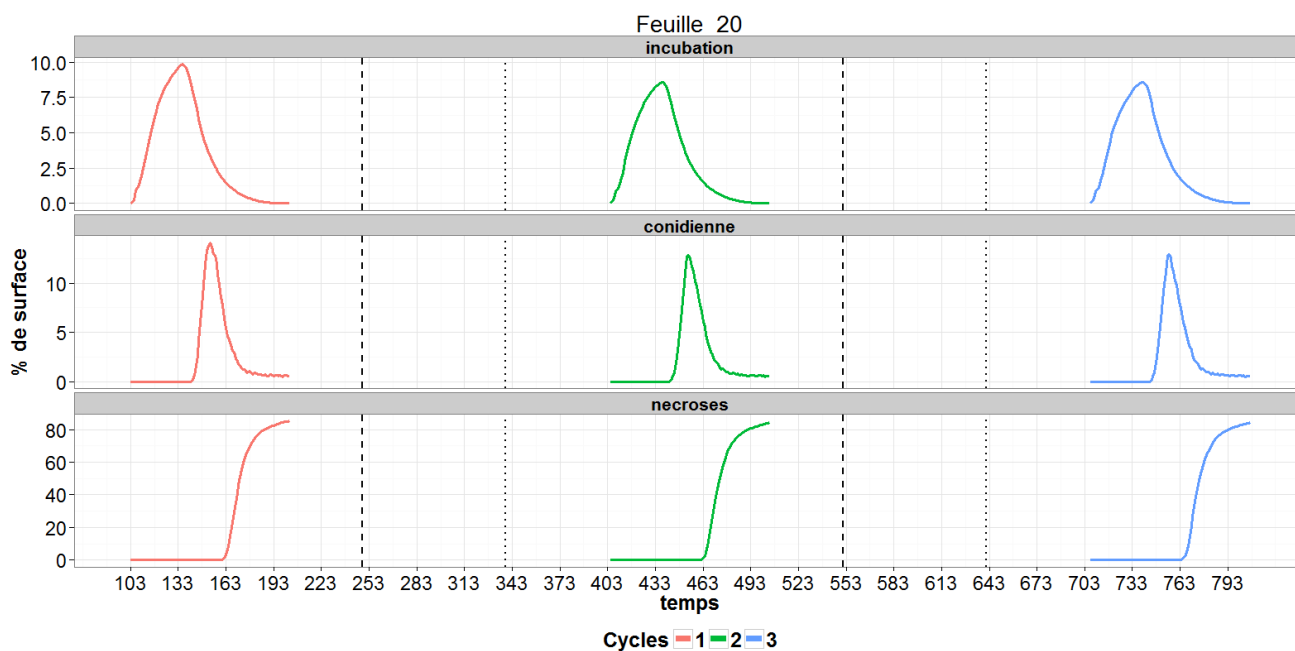


FIGURE 2.7 – Dynamique des pourcentages de surface malade de trois types de lésions (incubation, sporulantes en conidies et nécrosées) au cours de trois cycles de culture initiée avec 0.2 spores/mm²/jour sur la 20^{ème} feuille émise. Les barres verticales correspondent aux dates de floraison (tirets) et aux dates de récolte (pointillés).

2.2.3 Dynamique des inoculum au cours des cycles de culture

Dynamique de l'inoculum secondaire. L'inoculum secondaire est produit à deux étapes du cycle épidémique du pathogène : en début de cycle, à l'étape de reproduction asexuée (émission de conidies) et en fin de cycle, à l'étape de reproduction sexuée (émission d'ascospores). A l'échelle de la plante, l'inoculum secondaire est produit tout au long du cycle phénologique (Figure 2.8). Les quantités de spores sortantes augmentent jusqu'à une date intermédiaire entre la floraison et la récolte. L'inoculum secondaire asexué diminue plus tôt que l'inoculum sexué car sa production dans le cycle épidémique est plus précoce et qu'entre ces deux dates, très peu de nouvelles infections sont possibles (Figure 2.6).

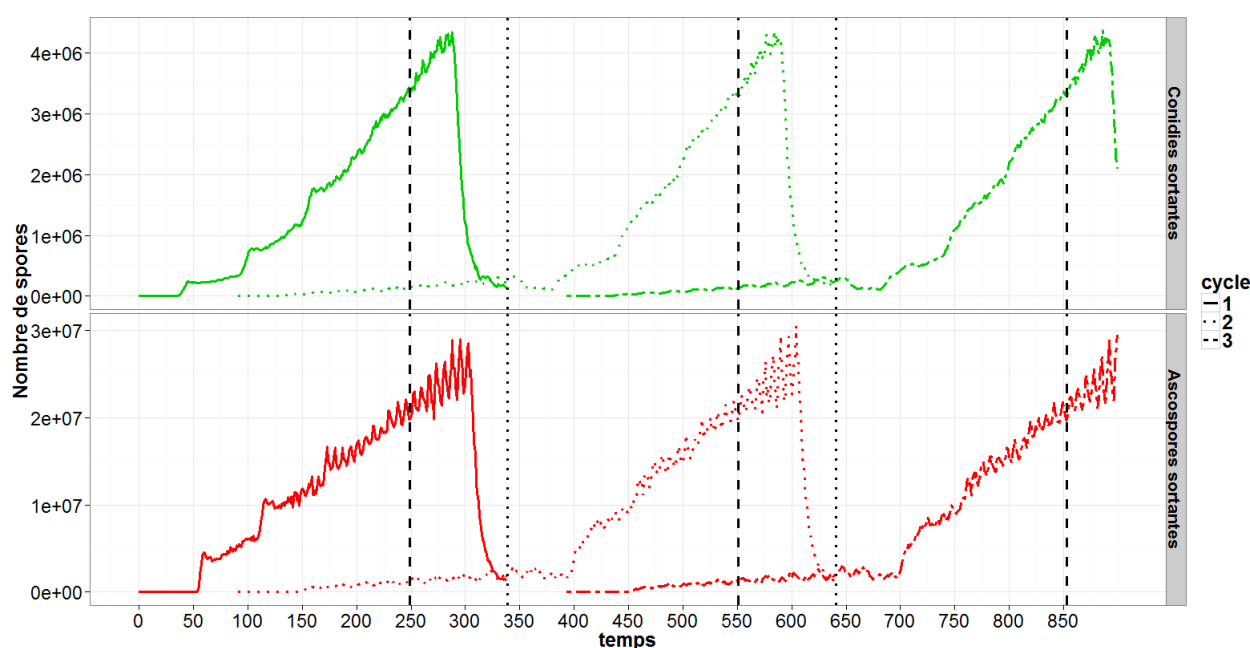


FIGURE 2.8 – Dynamique du nombre de conidies et d'ascospores sortantes au cours de trois cycles de culture initiée avec $0.2 \text{ spores/mm}^2/\text{jour}$. Les barres verticales correspondent aux dates de floraison (tirets) et aux dates de récolte (pointillés).

Dynamique de l'auto-inoculum et de l'allo-inoculum. L'auto-inoculum se définit par le nombre de spores émises par la plante et participant à son auto-infection. L'allo-inoculum se compose de spores provenant de l'extérieur (plantes voisines par exemple) et participe aussi à l'infection de la plante.

L'hypothèse selon laquelle l'efficacité d'infection des conidies et des ascospores est la même nous permet d'initier l'épidémie dans le modèle par le biais d'un seul type de spores. Ceci permet d'évaluer l'auto-infection de la plante. En effet, il est alors possible d'évaluer la quantité de spores émises par la plante et reçues d'elle même sachant qu'aucune spore de ce type ne fait partie de l'inoculum primaire.

En fixant un inoculum primaire composé uniquement d'ascospores, il nous est possible d'évaluer la quantité de conidies émises par la plante et celle participant à son auto-infection (Fi-

gure 2.9). Les différences observées entre les spores sortantes et entrantes sur la plante nous renseignent sur la part de d'auto-infection des conidies. Plus précisément, le ratio entre le nombre de conidies émises par la plante et le nombre de conidies reçues diminue dans le temps : en début de cycle, la quantité de spores sortantes est 20 fois supérieure à celle de spores entrantes, en fin de cycle ce ratio atteint 0,8.

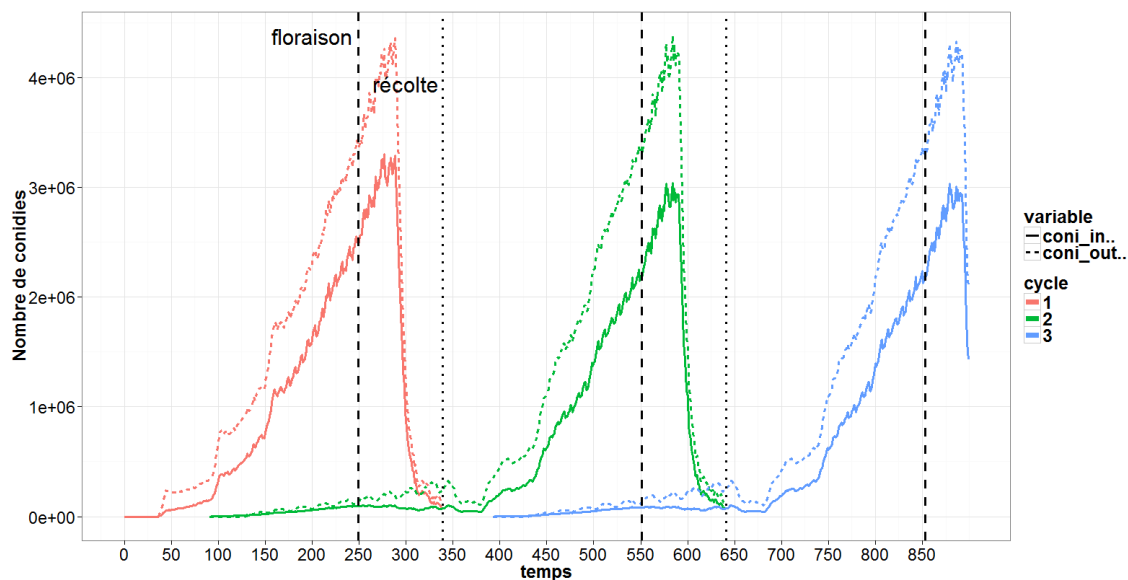


FIGURE 2.9 – Dynamique du nombre de conidies entrantes et sortantes au cours de trois cycles de culture initiée avec 0.2 spores/mm²/jour. Les barres verticales correspondent aux dates de floraison (tirets) et aux dates de récolte (pointillés).

Globalement, en étudiant plus en détail le ratio entre l'allo-inoculum et l'auto-inoculum (Figure 2.10), il est possible d'évaluer à quelles dates du cycle phénologique de la plante la quantité de spores provenant de l'extérieur devient moins importante que celle émise par la plante.

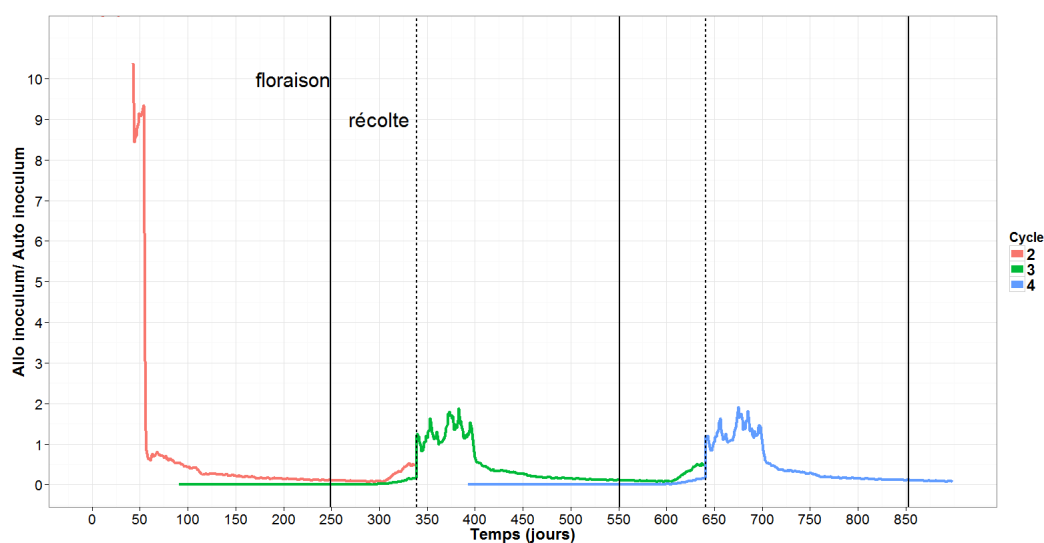


FIGURE 2.10 – Ratio allo-inoculum/auto-inoculum au cours des trois premiers cycles de culture initiée avec 0.2 spores/mm²/jour. Les barres verticales correspondent respectivement aux dates de floraison et aux dates de récolte.

2.2.4 Effet de l'inoculum primaire sur la vitesse du cycle infectieux

Un indicateur communément utilisé au champ pour connaître la vitesse du cycle infectieux est l'observation du rang de la plus jeune feuille présentant le symptôme recherché. Les feuilles étant ordonnées de la plus jeune à la plus ancienne (Figure 2.3), l'observation d'un symptôme sur la plus jeune feuille sachant le rythme d'émission foliaire de la plante permet d'évaluer la vitesse du cycle infectieux. Les deux principaux symptômes recherchés sont les lésions conidiennes dites "matures", l'indice est noté PJFM (plus jeune feuille ayant une lésion conidienne), et les nécroses, l'indice est noté PJFN (plus jeune feuille nécrosée) [Carlier et al., 2000]. Le modèle permet de simuler ces indices (Figure 2.11) et d'évaluer par exemple l'effet de l'inoculum primaire sur la variabilité de ces indices.

Dans l'exemple présenté, l'inoculum primaire ne semble pas avoir d'impact sur le rang des PJFN et PJFM.

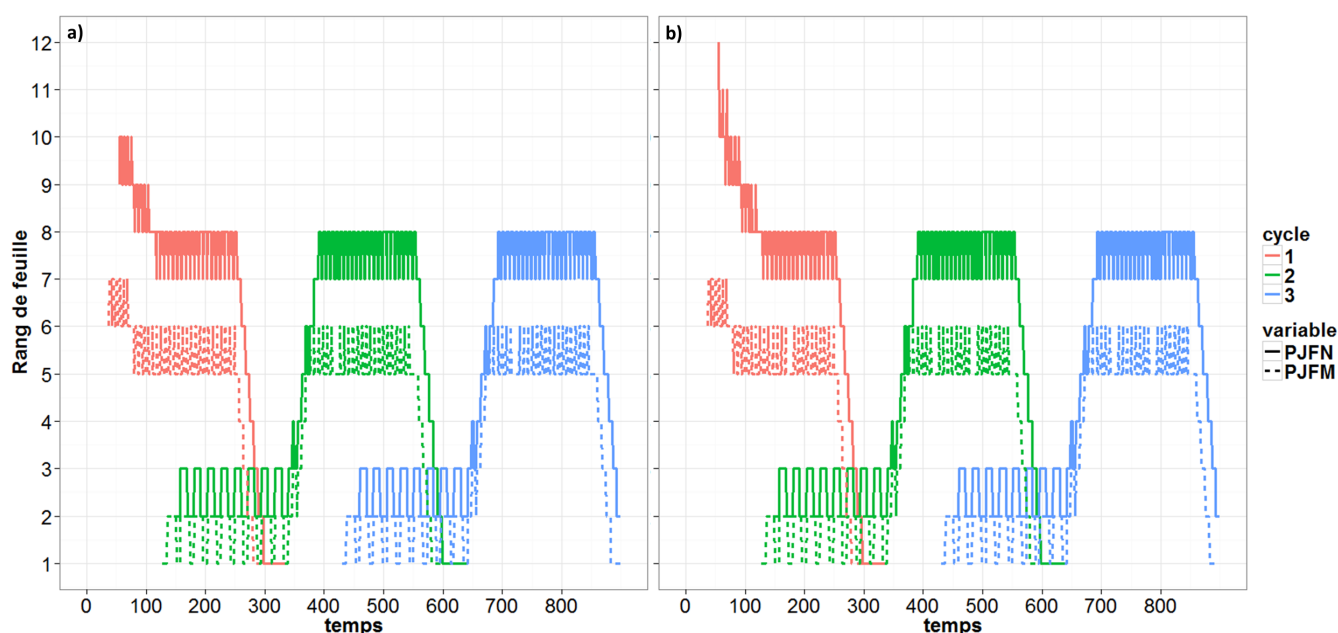


FIGURE 2.11 – Evaluation des PJFN et PJFM selon deux inoculum primaires contrastés a) 0.2 spores/mm²/jour et b) 0.01 spores/mm²/jours.

2.2.5 Effet de la résistance du bananier sur la dynamique parasitaire

Considérons une variété S sensible comme définie dans le tableau 2.2 et une variété R résistante qui a les mêmes paramètres épidémiologiques que la variété sensible excepté la vitesse d'extension des lésions $lgr = 0.08$ mm²/jour (soit environ deux fois moins rapide que la variété S décrite) et la capacité de sporulation sexuée $sc_s = 58$ spores/mm² de nécrose (soit environ 3 fois moins que la variété S décrite). Les sorties du modèle permettent d'étudier la variabilité des surfaces de sept types de lésions : les lésions en incubation, les lésions immatures conidiennes (de tailles insuffisantes pour l'émission de spores asexuée), les lésions sporulantes conidiennes, les lé-

sions immatures ascosporees (lésions non coalescentes), les lésions sporulantes ascosporees et lésions inactives (non sporulantes). Nous nous intéressons par la suite aux quatre types de lésions : incubation, sporulantes conidiennes, sporulantes ascosporees ou inactives pour les deux variétés décrites précédemment, sur chacune des feuilles de la plante à une date donnée (Tableau 2.3).

Feuille	Surfaces (mm ²)							
	lésions en incubation		conidienne		nécroses sporulantes et non sporulantes		totale malade	
	S	R	S	R	S	R	S	R
F1	62868	15773	0	0	0	0	62868	15773
F2	125914	32432	0	0	0	0	125914	32432
F3	157139	41440	0	0	0	0	157139	41440
F4	169285	45620	0	0	0	0	169285	45620
F5	147170	40703	0	0	0	0	176433	47712.1
F6	80002	22894	86648.2	2543.48	0	0	249178	57418.9
F7	40143	11752	185291	40807.1	0	0	412420	75721.4
F8	21310	6528	73667.9	34338.5	3362.74	0	608903	114128
F9	10592	3597	24715.3	18203.1	308084	8032.6	727376	171442
F10	4730	1957	9127.35	9926.16	548716	90134.7	738500	231844
F11	1697	965	5952.3	5104.18	609788	162531	716116	255559
F12	504	484	4729.25	3016.26	619291	210816	694449	274229

TABLEAU 2.3 – Évaluation des surfaces de trois types de lésions et de la surface malade totale pour deux variétés (sensible S et résistante R). L'évaluation est faite sur les 12 feuilles de la plante, au milieu de son premier cycle phénologique : 30 jours avant la floraison.

Les résultats montrent que la surface des lésions en incubation et sporulantes en conidies décroît avec le rang de feuille alors qu'elle augmente pour les surfaces nécrosées. Il est à noter que deux rangs de feuilles (F6 et F7) n'ont que des lésions sporulantes conidiennes. Les simulations numériques montrent que sur la variété résistante, la surface malade est inférieure à celle de la variété sensible sur tous les rangs de feuilles, et ce du fait d'une surface nécrosée plus tardive et plus petite.

2.3 Discussion

Le modèle SiBaToka développé est le premier modèle simulant la dynamique du pathogène *Mycosphaerella fijiensis* couplée à la croissance d'un bananier sur plusieurs cycles de culture.

Le modèle permet de mieux comprendre la dynamique épidémique grâce à la flexibilité de son

paramétrage et aux sorties disponibles à différentes échelles spatiales (plante, feuilles, compartiments de feuilles). En effet, tous les paramètres d'entrée dépendent de caractéristiques variétales. Il est donc possible d'explorer les sorties du modèle en faisant varier les paramètres épidémiologiques du pathogène et les paramètres phénologiques et architecturaux de la plante. Le modèle informatique a été conçu de manière à ce que les nombres de spores, toutes les surfaces et les nombres de lésions, dans les différents états épidémiologiques (lésions en incubation, lésions productrices de spores asexuées, lésions latentes, nécroses sporulantes et non sporulantes) soient accessibles dans chacun des compartiments de feuilles présents sur la plante au cours du temps.

Les méthodes de lutte peuvent être évaluées grâce au modèle. Par exemple, afin de mieux comprendre l'effet de la croissance du bananier sur les niveaux épidémiques, nous avons mis en évidence l'effet du rythme d'émission foliaire sur le pourcentage de nécrose présent sur une plante au cours du temps. Il est donc possible de proposer des pratiques culturales visant à augmenter le rythme d'émission foliaire et de fait, à réduire la quantité de maladie sur la plante. Le modèle peut permettre aussi d'aider à la conception de nouvelles méthodes de lutte. Par exemple, en évaluant la part de l'auto-infection au cours du cycle, il est possible de proposer, à des dates clés, de réduire l'auto-infection par le biais d'effeuillages.

L'inoculum primaire utilisé pour initier l'épidémie est difficile à quantifier et son évaluation au champ est compliquée [van Maanen and Xu, 2003]. Nous avons montré que cette quantité n'a pas d'influence sur le rang des PJFN et PJFM simulés, cependant il serait intéressant de pouvoir évaluer l'impact de l'inoculum primaire sur les sorties du modèle.

Dans le modèle, la croissance des lésions est décrite par une fonction logistique. Dans sa définition, la taille maximale d'une lésion est le compartiment, or il est admis qu'une lésion isolée possède une surface maximale [Jones, 2000]. On pourrait alors proposer une expérimentation permettant de mesurer cette surface afin d'intégrer cette connaissance dans le modèle. Le paramètre de cette fonction lgr (vitesse d'extension des lésions) est cependant fixe dans le temps, alors que Fouré [1982] décrit une accélération du cycle épidémique lorsque la densité de lésion est élevée. Une amélioration du modèle pourrait être de simuler la croissance des lésions en intégrant un paramètre de densité-dépendance.

Dans le modèle, les lésions émettent une quantité fixée de spores par mm^2 et par jour, cependant la quantité de spores produite par une lésion nécrosée varie dans le temps [Jones, 2000]. On pourrait alors proposer de prendre cela en compte comme par exemple Eikemo et al. [2011] qui évaluent six fonctions décrivant la production d'ascospores d'un pathogène du poirier.

La dispersion des spores est décrite par deux fonctions de dispersion exponentielles. Il serait intéressant de comparer plusieurs fonctions de dispersion afin de trouver pour chacun des types de spores une fonction adéquate comme le proposent Rieux et al. [2014] qui comparent plusieurs fonctions de dispersion des conidies et ascospores sur une distance d'un kilomètre. L'amélioration du choix de la fonction de dispersion peut permettre d'améliorer les prédictions du modèle.

Plus généralement, il serait intéressant de prendre en compte le fait que la croissance du ba-

nanier dépend de l'état sanitaire de la plante, comme le suggèrent [Mobambo et al. \[1993\]](#).

Ce modèle pourra être adapté à de plus grandes échelles spatiales, en effet, mieux comprendre la dynamique de l'inoculum entre plants pourrait permettre de développer des méthodes de lutte efficaces par agencements variétaux comme le propose par exemple [Papaix \[2011\]](#).

Chapitre 3

Analyse de sensibilité du modèle SiBaToka

Sommaire

3.1 Introduction	85
3.2 Matériel et méthodes	86
3.2.1 Le modèle SiBaToka	86
3.2.2 Méthodes d'analyse de sensibilité	92
Méthode de Morris	92
Analyse eFAST	93
Démarche suivie pour l'analyse de sensibilité du modèle	94
Gammes de variation des paramètres	95
3.3 Résultats	96
3.3.1 Évaluation de l'effet des composantes de résistance de la plante et de la dis- persion de l'inoculum secondaire	96
Par la méthode de Morris	96
Par la méthode e-FAST	98
3.3.2 Évaluation de l'effet des caractères agronomiques et génétiques de la plante	102
Par la méthode de Morris	102
Par la méthode e-FAST	103
3.3.3 Évaluation du modèle complet	105
Par la méthode de Morris	105
Par la méthode e-FAST	106
3.4 Discussion	108
3.4.1 Discussion des principaux résultats	108
Trois composantes de résistance de la plante aux influences majeures	108
Inoculum primaire et inoculum secondaire	109
Caractéristiques agronomiques et génétiques de la plante	110
3.4.2 Discussion des méthodes et du modèle	110

Méthodes utilisées et perspectives	110
Sensibilité des sorties d'intérêt à la variation des paramètres : conséquences pour l'amélioration du modèle	111

Dans ce chapitre, nous évaluons l'effet des paramètres du modèle SiBaToka présenté précédemment en utilisant deux méthodes d'analyses de sensibilité. Ce chapitre est construit sous forme d'article : nous prévoyons sa publication dans la revue *Plant Pathology*.

3.1 Introduction

Les modèles de cultures intégrant divers pathogènes permettent d'améliorer la compréhension des processus épidémiologiques et plus précisément d'évaluer le fonctionnement de l'interaction entre la dynamique de la culture et celle du pathogène (Gilligan [2008], De Wolf and Isard [2007]). Ces modèles sont des outils majeurs d'aide à la décision pour le contrôle des maladies en permettant l'amélioration de la gestion des cultures [Jeger, 2004].

Décrire les processus mis en jeu dans le développement d'un pathogène sur une culture nécessite généralement l'intégration de beaucoup de paramètres d'entrée. Par exemple Richter et al. [2010] décrivent un modèle de culture du blé par le biais de 91 paramètres ou encore Lô Pelzer et al. [2010a] analysent un modèle simulant le phoma du colza décrit avec 49 paramètres d'entrée. L'incertitude liée aux valeurs données à ces nombreux paramètres entraîne une incertitude sur les prédictions du modèle [Lamboni et al., 2009]. Afin de calibrer efficacement un modèle, il est nécessaire de connaître les paramètres les plus influents sur les sorties du modèle pour pouvoir mettre en place des expérimentations permettant d'affiner leurs valeurs [Wang et al., 2013].

L'analyse de sensibilité globale d'un modèle est une approche permettant d'évaluer l'effet de la variabilité des paramètres d'entrée sur des sorties d'intérêt. Elle permet d'identifier les paramètres clés i.e. les paramètres qui ont la plus grande influence sur les sorties du modèle, de quantifier leurs importances et par là même d'explorer la structure du modèle [Faivre et al., 2013]. L'identification de ces paramètres peut permettre alors un meilleur contrôle de la maladie : améliorer les connaissances sur ces paramètres augmente la précision du modèle et identifier ces paramètres peut permettre de proposer des mesures de contrôle efficaces (Casadebaig et al. [2012], DeJonge et al. [2012], Confalonieri et al. [2010b]).

A l'échelle d'une plante, un modèle épidémiologique de simulation de la cercosporiose noire du bananier, considérée comme la maladie foliaire la plus grave de cette culture [Jones, 2000], a été développé en temps discret et simule le cycle épidémique de la maladie et la croissance du bananier pendant plusieurs cycles [Landry et al., in prep]. Ce modèle simule, suivant une quantité donnée d'inoculum primaire (nombre de spores dans l'air se déposant par mm^2 de feuille et créant des lésions), la croissance et la sporulation des lésions créées. De ce fait, ce modèle permet d'étudier l'auto-inoculum (spores provenant de la plante) permettant la réinfestation de la plante par elle-même. En effet, ce modèle considère que les deux types de spores (asexuées et sexuées) produites par le pathogène responsable de la cercosporiose noire se dispersent sur les feuilles de la même plante pour créer de nouvelles lésions.

Vis-à-vis de ce pathosystème, il a été suggéré l'existence deux types de résistance des bananiers : une résistance totale et une résistance partielle [Jones, 2000]. Contrairement à la résistance totale, dans le cas de la résistance partielle, le pathogène peut réaliser toutes les étapes de son cycle infectieux (infection, sporulation, dispersion) mais généralement de façon moins rapide. Il a été montré que les composantes de résistance partielle des bananiers agissent sur les principales

étapes du cycle infectieux de *Mycosphaerella fijiensis* (Abadie et al. [2003], Abadie et al. [2009]). Ainsi, les composantes de résistance de bananiers sensibles correspondent à des valeurs de paramètres différentes de bananiers sensibles à la cercosporiose noire.

La non-linéarité des processus qui interviennent dans le modèle et ses sorties temporelles suggèrent que la contribution de chaque paramètre à la variation des sorties peut aussi varier dans le temps [Song et al., 2013]. En effet, certains paramètres peuvent être influents au début des différents processus et perdre de l'influence dans les étapes suivantes comme le décrivent Song et al. [2012] sur un modèle de croissance d'une forêt.

Le but est d'explorer le comportement de SiBaToka à la variation de ses paramètres d'entrée (correspondant à une large gamme de sensibilité variétale à la maladie) et d'identifier les paramètres les plus influents. Nous avons choisi des situations de pressions parasitaires contrastées de façon à représenter la variabilité de l'inoculum primaire existante et d'étudier son effet sur les paramètres du modèle. Ce dernier comportant deux modules, nous avons d'abord étudié la variabilité des paramètres du module décrivant le cycle épidémique du pathogène *Mycosphaerella fijiensis* ("epidemic submodel" du Chapitre 2) en considérant une croissance optimale de la plante, puis nous avons étudié l'influence des paramètres du module plante ("plant submodel" du Chapitre 2) en fixant des paramètres moyens pour le cycle épidémique et pour finir nous avons étudié l'influence des paramètres du modèle complet SiBaToka.

Dans un premier temps, nous présenterons tout d'abord brièvement le modèle SiBaToka et ses deux modules ainsi que les méthodes d'analyse de sensibilité globale utilisées, puis nous détaillerons les résultats obtenus pour chacun des modules et pour le modèle complet. Pour finir, nous discuterons de ces résultats en comparant d'une part les méthodes utilisées et d'autre part les sorties en fonction des différentes doses d'inoculum primaires.

3.2 Matériel et méthodes

3.2.1 Le modèle SiBaToka

Le modèle SiBaToka a été décrit en détail dans [Landry et al., in prep]. Il est composé de deux modules, un premier module décrivant la croissance successive d'un bananier et de ses rejets et un module épidémiologique décrivant les étapes du cycle épidémique de *Mycosphaerella fijiensis* (Figure 3.1).

Ce modèle est défini en temps discret (pas de temps de un jour) et est basé sur l'hypothèse que les conditions environnementales et climatiques sont prises en compte dans les paramètres d'entrée du modèle. Il intègre 31 paramètres d'entrée : 11 liés au cycle épidémique du pathogène et 20 liés à la croissance de la plante sur trois cycles de culture.

Le module dynamique de croissance de la plante. Il décrit spatialement et temporellement la croissance d'un bananier sur plusieurs cycles de culture.

Il simule la croissance d'un vitroplant depuis la plantation jusqu'à la récolte (environ 12 mois) et celle de ses rejets successifs, décrivant le développement foliaire de chaque plante.

Un bananier possède dans son pseudo-tronc un nombre prédéfini de feuilles enroulées qui se déploient une à une (plus ou moins rapidement avec une moyenne d'une feuille par semaine) jusqu'à la floraison. De la floraison à la récolte du régime, le bananier n'émet plus de feuilles.

Les feuilles sur chaque plante sont discrétisées en sept compartiments correspondant aux sept portions de feuille qui sont déployées successivement pendant le déroulement de chacune des feuilles.

Le module épidémiologique. Il décrit les trois étapes du cycle infectieux du pathogène dans chacun des compartiments de feuilles.

L'infection (première étape) est modélisée par le nombre de lésions en incubation résultant de l'inoculation d'une quantité de spores provenant de l'inoculum primaire et secondaire produit par la plante et de l'efficacité d'infection. Après une durée d'incubation, les lésions apparaissent et mesurent 1 mm^2 (surface minimale d'une lésion).

La croissance des lésions dans chaque compartiment est simulée de manière logistique.

La deuxième étape simule la sporulation du champignon. La sporulation asexuée a lieu lorsque les lésions ont atteint une taille minimale (qui correspond à la période de latence asexuée) et dure un temps défini (durée infectieuse conidienne).

La dispersion (troisième étape) des spores asexuées produites (conidies) est modélisée par une fonction de dispersion exponentielle dont le paramétrage ne doit pas permettre d'événement à longue distance, ces spores étant principalement déplacées par le ruissellement de l'eau sur les feuilles.

Une fois la sporulation asexuée terminée, les lésions continuent de croître et en fonction de la densité de lésions par compartiment, les lésions peuvent coalescer formant des nécroses et émettre des spores sexuées. La rencontre des lésions est modélisée de façon stochastique par une probabilité de rencontre entre deux lésions de types sexuels compatibles. Les spores sexuées émises (ascospores) peuvent se propager à longue distance. Nous avons modélisé leur dispersion entre compartiments foliaires par une fonction exponentielle.

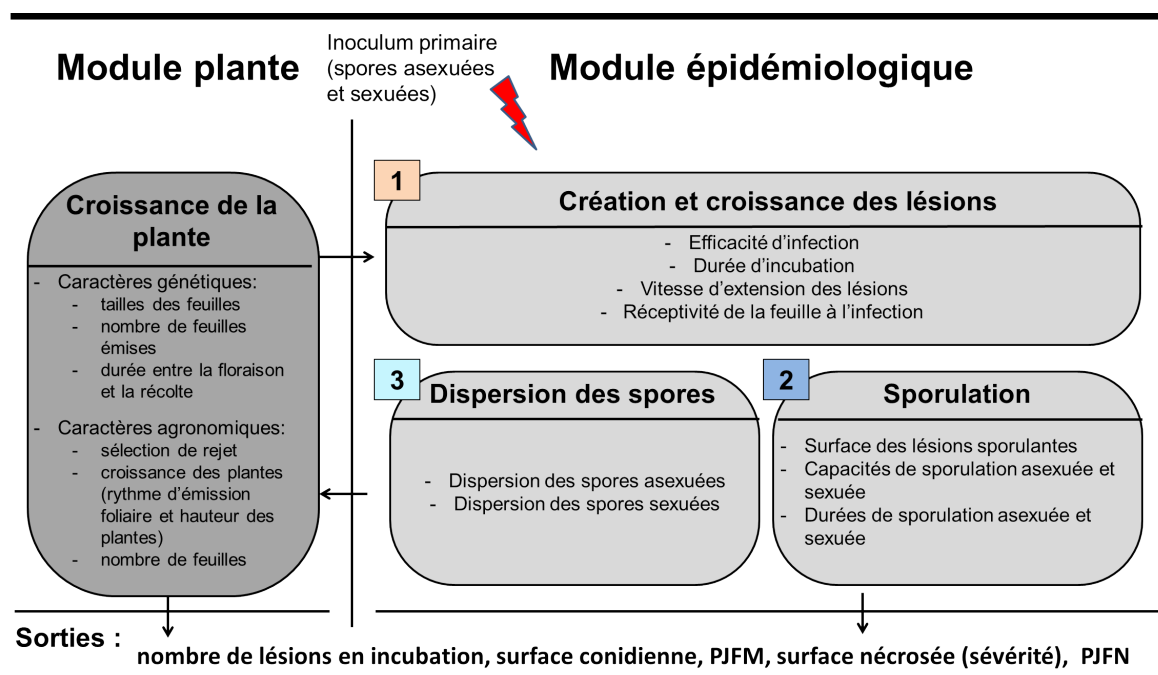


FIGURE 3.1 – Représentation schématique du modèle SiBaToka composé d'un module de croissance de la plante (module plante) et d'un module épidémiologique

Variables d'entrée et variables de sortie. Les 20 paramètres d'entrée du module décrivant la croissance de la plante peuvent être classés en deux groupes : les paramètres liés aux caractères agronomiques et ceux liés aux caractères génétiques spécifiques à chaque type de cultivar (Tableau 3.1). Les 11 paramètres décrivant le cycle épidémique du pathogène sont classés en trois groupes, ceux liés à la création et la croissance des lésions, ceux liés aux sporulations asexuée et sexuée et ceux liés à la dispersion des spores produites par les lésions (Tableau 3.2)

Les variables de sortie sont des variables épidémiologiques : des nombres de lésions en incubation, des surfaces de lésions sporulantes en conidies (surfaces conidiennes), des surfaces nécrosées (sévérité de la maladie), des surfaces malades (la surface non-saine de la plante) et des rangs de feuilles. Les feuilles d'un bananier sont numérotées par leur hauteur sur la plante, la feuille de rang 1 est la feuille la plus haute soit la dernière émise (la plus jeune sur le bananier). Deux paramètres semi-quantitatifs sont communément utilisés pour évaluer l'état d'infestation d'une parcelle ou d'un bananier : le rang de la plus jeune feuille ayant une lésion mature (sporulante en conidies) noté PJFM, et le rang de la plus jeune feuille ayant une nécrose (lésion sporulante en ascospores), noté PJFN [Carlier et al., 2002]. Ces évaluations sont aussi des variables de sortie du modèle.

	Paramètre	Description	Valeur nominale	Intervalle	Unité	Référence
Caractères génétiques	1 : Δ_{fh}^{**}	temps entre la floraison et la récolte	84	[59, 139]	jours	Irish et al. [2013]
	2 : ϕ_L^V	angle entre deux feuilles consécutives (vertical)	10	[7, 11]	degrés	Stover and Simmonds [1987]
	3 : L_L^{max*}	longueur de feuille maximale	2016	[1925, 2520]	mm	Stover and Simmonds [1987] [Landry et al., in prep]
	4 : L_L^{min*}	longueur de feuille minimale	486	[400, 630]	mm	[Landry et al., in prep]
	5 : W_L^{max}	largeur de feuille maximale	846	[810, 930]	mm	Stover and Simmonds [1987]
	6 : W_L^{min*}	largeur de feuille minimale	253	[215, 320]	mm	[Landry et al., in prep]
	7 : ϕ_L^H	angle phyllotaxique	156	[140, 180]	degrés	Lassoudière [2012]
	8 : l_p^0	rapport longueur pétiole/feuille à la plantation	0.162	[0.1012, 0.235]	-	[Landry et al., in prep]
	9 : l_p^1	rapport longueur pétiole/feuille à l'apparition du rejet	0.238	[0.13, 0.3913]	-	[Landry et al., in prep]
	10 : H_L^0	hauteur du point d'insertion de la feuille la plus basse à la plantation	360	[280, 490]	mm	[Landry et al., in prep]
	11 : $N_L^{flower**}$	nombre de feuilles émises avant la floraison	36	[25, 50]	-	Robinson [2003]
caractères agronomiques	12 : H_L^1	hauteur du point d'intersection de la feuille la plus basse à l'apparition du rejet	449	[325, 570]	mm	[Landry et al., in prep]
	13 : N_L^0*	nombre de feuilles émises à l'apparition du rejet sélectionné	18	[13, 25]	-	Lassoudière [2006]

14 : l_{node}	longueur des entrenœuds	87	[25, 150]	mm	Robinson [2003]
15 : ler^{init}	rythme d'émission foliaire initial	1.32	[1.368, 2.02]	feuille /se- maine	Stover and Sim- monds [1987]
16 : $ler^{final} **$	rythme d'emission foliaire final	1	[0.7, 1.33]	feuille /se- maine	Stover and Sim- monds [1987]
17 : $N_L^{harvest}$	nombre de feuilles sur la plante à la récolte	5	[1, 7]	-	Irish et al. [2013]
18 : $N_L^{max} **$	nombre de feuilles maximal sur la plante	12	[8, 16]	-	Irish et al. [2013]
19 : $N_L^{0,init}$	nombre de feuilles sur la plante à la plantation	4	[3, 6]	-	[Landry et al., in prep]
20 : $N_L^{1,init} **$	nombre de feuilles du rejet à la récolte du pied mère	9	[6, 11]	-	Lassoudière [2012]

TABLEAU 3.1 – Paramètres du module plante soumis aux analyses de sensibilité. Les astérisques * indiquent les paramètres conservés à la suite de l'analyse selon la méthode de Morris et soumis à l'analyse selon la méthode e-FAST pour le module plante, ** désigne les paramètres soumis à l'analyse e-FAST pour le module plante et le modèle SiBaToka complet.

	Paramètre	Description	Valeur nominale	Intervalle	Unité	Référence
Création et croissance des lésions	1 : ei^{**}	efficacité d'infection sur une feuille nouvellement émise (nombre de lésions créées par une spore)	0.0052359	[0.0004, 0.01853]	-	[Landry et al., in prep]
	2 : ip^{**}	durée d'incubation : temps entre l'infection et l'apparition de la lésion	23	[11, 36]	jours	Fouré [1982]
	3 : lgr^{**}	taux de croissance des lésions	0.15	[0.01, 0.33]	mm ² /jour	[Landry et al., in prep]
	4 : $ei^{age **}$	effet de l'âge de la feuille sur l'efficacité d'infection	0.06	[0.03, 0.09]	-	-
Sporulation	5 : $a^{min **}$	taille minimale d'une lésion capable d'émettre des spores asexuées	2	[1.5, 2.5]	mm ²	Fouré [1982]
	6 : sc_a^{**}	capacité de sporulation asexuée	7	[1, 60]	nombre /mm ²	Fouré [1982]
	7 : sc_s^{**}	capacité de sporulation sexuée	182	[58, 436]	nombre /mm ²	Fouré et al. [1984] Burt et al. [1999]
	8 : Δ_a^{**}	durée de sporulation asexuée	8	[3, 18]	jours	[Landry et al., in prep]
	9 : Δ_s^{**}	durée de sporulation sexuée	1	[1, 3]	jours	Jones [2000] Gauhl [1994]
Dispersion	10 : β_a^{**}	paramètre de décroissance de la fonction de dispersion des spores asexuées	0.001	[0.0001, 0.01]		Rieux et al. [2013]
	11 : β_s^*	paramètre de décroissance de la fonction de dispersion des spores sexuées	0.00001	[0.000001, 0.0001]		Rieux et al. [2013]

TABEAU 3.2 – Paramètres du module épidémiologique soumis aux analyses de sensibilité. Les astérisques * indiquent les paramètres conservés à la suite de l'analyse selon la méthode de Morris et soumis à l'analyse selon la méthode e-FAST pour le module épidémiologique, ** désigne les paramètres soumis à l'analyse e-FAST pour le module épidémiologique et le modèle SiBaToka complet.

3.2.2 Méthodes d'analyse de sensibilité

De nombreuses méthodes d'analyse de sensibilité globale de modèles numériques existent dans la littérature, une revue de ces méthodes est détaillée dans Iooss [2010] et plus généralement dans Faivre et al. [2013].

La méthodologie qui consiste à utiliser d'abord une méthode qualitative permettant la sélection de paramètres puis une méthode quantitative a été appliquée dans l'analyse de modèles décrivant différents écosystèmes, par exemple des écosystèmes marins (Ciric et al. [2012], Morris et al. [2014]), des écosystèmes forestiers [Song et al., 2012], ou encore des agro-écosystèmes (DeJonge et al. [2012], Confalonieri et al. [2010b]).

En suivant cette méthodologie, nous avons tout d'abord effectué une première analyse de sensibilité par la méthode de "screening" de Morris [Morris, 1991]. Cette méthode permet d'isoler rapidement les paramètres pouvant être négligés à partir d'une analyse qualitative de la sensibilité d'une variable de sortie. Cette méthode est largement utilisée lorsque les paramètres sont nombreux et les temps de calculs longs [Richter et al., 2010] car elle propose des résultats robustes avec peu de simulations [Confalonieri et al., 2010a] [Campolongo et al., 2007].

Nous avons ensuite utilisé la méthode extended-FAST (Fourier Amplitude Sensitivity Test) [Saltelli et al., 1999] basée sur la décomposition de la variance fonctionnelle qui permet le calcul des indices de Sobol de premier ordre (indices principaux) et des indices totaux et donc de quantifier l'impact des paramètres sur les sorties d'intérêt. Cette méthode est considérée comme la plus efficace en terme de nombre d'évaluations du modèle (Saltelli et al. [1999], Makowski et al. [2006b]) et est utilisée pour l'évaluation de nombreux modèles de cultures (Makowski et al. [2006b], Varella et al. [2010], Zhao et al. [2014], Wang et al. [2013], Drouet et al. [2011]).

Méthode de Morris

La méthode de Morris utilise une discrétisation régulière de l'espace des paramètres (grille multi-dimensionnelle) dans laquelle l'influence de chaque paramètre d'entrée est évaluée en comparant des simulations entre lesquelles seul un paramètre d'entrée varie. En effet, la méthode de Morris fait partie des méthodes utilisant un plan d'expérience OAT "One-At-a-Time" qui consiste à changer le niveau d'un paramètre à la fois, en utilisant un nombre déterminé de niveaux par paramètre. Dans cette méthode, un plan d'expérience OAT est répété r fois aléatoirement dans l'espace des paramètres.

Chaque répétition i ($i = 1, \dots, r$) permet d'évaluer un effet élémentaire du paramètre $j : E_j^i$ (accroissement de la sortie du modèle entre deux valeurs consécutives) et les r trajectoires permettent le calcul de r effets élémentaires par paramètre en $N = r(K + 1)$ simulations où K est le nombre de paramètres d'entrée à évaluer.

Pour chaque variable d'entrée une distribution de valeurs représentant son effet dans les différentes répétitions est obtenue. Ces distributions sont caractérisées par la moyenne (μ_j^*) et l'écart-

type (σ_j) des effets élémentaires E_j^i , avec

$$\mu_j^* = \frac{\sum_{i=1}^r |E_j^i|}{r}$$

Un facteur possédant une forte moyenne μ_n^* est un facteur ayant une forte influence sur la sortie du modèle, une forte valeur d'écart-type σ_j indique soit l'existence d'une interaction entre le paramètre j et un autre, soit l'existence d'un effet non linéaire de ce paramètre. Les détails supplémentaires de cette méthode sont donnés dans le [Chapitre 1](#).

Le calcul de ces effets élémentaires a été réalisé à l'aide du logiciel R [[R Development Core Team, 2008](#)] et du package Sensitivity (Gilles Pujol). Nous avons utilisé cette méthode pour les deux modules et le modèle complet SiBaToka en choisissant une discrétisation de l'intervalle d'entrée de chaque paramètre en 6 niveaux (Tableaux [3.1](#) et [3.2](#)), 1000 répétitions pour le module épidémiologique et 500 répétitions pour le module plante et pour le modèle complet compte tenu des temps de calculs (Tableau [3.3](#)).

Module	Paramètres	Niveaux	Répétitions	Évaluations du modèle	Temps de calcul
Épidémiologique	11	6	1000	12000	56953 sec.
Plante	20	6	500	10500	108904 sec.
SiBaToka	31	6	500	16000	136585 sec.

TABLEAU 3.3 – Détail de l'utilisation de la méthode de Morris pour les deux modules et le modèle complet SiBaToka

Analyse eFAST

Cette méthode consiste à déterminer quelle part de la variance de la sortie d'intérêt est due à la variance de chaque paramètre. Elle fournit la part de variance expliquée par les variations d'un paramètre j indépendamment de la valeur des autres paramètres (effet principal S_j) et la part de variance expliquée par j y compris à travers ses interactions avec les autres paramètres (effet total ST_j). Les indices sont évalués par un balayage de l'espace des paramètres d'entrée de manière à ce que chaque paramètre soit exploré selon une fréquence de nombres entiers choisis. Les détails de cette méthode sont présentés dans le [Chapitre 1](#) et disponibles dans [[Monod et al., 2006](#)]. Cette méthode ne repose sur aucune hypothèse concernant le modèle et peut donc être utilisée sur des modèles non linéaires ou non monotones. Elle présente l'inconvénient de ne pas calculer les termes d'interaction en détail.

Le calcul de l'effet principal S_j et de l'effet total ST_j de chaque paramètre j , sélectionné comme le plus influent par la méthode de Morris, a été réalisé à l'aide du logiciel R [[R Development Core Team, 2008](#)] et du package Sensitivity (Gilles Pujol). Chaque indice a été calculé, pour les deux modules et le modèle complet SiBaToka, avec le même nombre de simulations pour chaque paramètre (Tableau [3.4](#)). Les valeurs des paramètres ont été tirés dans des lois uniformes sur les

intervalles définis dans les tableaux 3.1 et 3.2. Afin d'optimiser les temps de calculs, nous avons utilisé une dizaine de cœurs par module en lançant les analyses en parallèle.

Module	Paramètres	Simulations	Évaluations du modèle
Épidémiologique	11	5000	55000
Plante	10	5000	50000
SiBaToka	15	5000	75000

TABLEAU 3.4 – Détails de la méthode extended-FAST pour les deux modules et le modèle complet SiBaToka

Démarche suivie pour l'analyse de sensibilité du modèle

Nous avons conduit l'analyse de sensibilité sur plusieurs sorties d'intérêt à plusieurs dates du cycle phénologique de la plante selon les modules analysés (Tableau 3.5), en considérant dans chacun des cas deux quantités différentes d'inoculum primaire. Ces quantités ont été choisies de manière à évaluer les possibles différences de pression parasitaire. La faible pression parasitaire est définie par une quantité journalière de conidies entrantes sur la plante de 0.0005 conidies/mm² et de 0.0045 ascospores/mm². La forte pression parasitaire se caractérise par une quantité journalière de conidies entrantes sur la plante de 0.01 conidies/mm² et de 0.19 ascospores/mm².

L'analyse de sensibilité du module épidémiologique permet d'étudier d'une part le poids épidémiologique des composantes de résistance et d'autre part de mieux comprendre le fonctionnement épidémique des cercosporioses (en particulier la production d'inoculum secondaire et sa dispersion). Nous avons donc construit les analyses en suivant les étapes du cycle épidémique du pathogène (Figure 3.1) :

1. la création et la croissance des lésions, où la sortie d'intérêt est le nombre de lésions en incubation sur la plante.
2. la sporulation asexuée où les sorties d'intérêt sont le pourcentage de surface conidienne sur la plante et le rang de la plus jeune feuille ayant une lésion conidienne (PJFM), traduisant la latence asexuée (temps entre la contamination et l'apparition de lésions sporulantes en conidies).
3. la sporulation sexuée où les sorties d'intérêt sont le pourcentage de nécroses sur la plante et le rang de la plus jeune feuille ayant une nécrose (PJFN), traduisant la latence sexuée (temps entre la contamination et l'apparition d'une nécrose sporulante en ascospore).

Plus globalement, nous étudierons aussi la sortie d'intérêt concernant le pourcentage de maladie (tous type de lésions) sur la plante. Les analyses sont réalisées à trois dates : la date d'initiation florale (début de remontée de la fleur dans le pseudotrunc correspondant approximativement au milieu du cycle phénologique de la plante), la date de la floraison (date d'arrêt de l'émission de

feuilles) et 30 jours avant la récolte (date importante pour l'impact de la maladie sur la production de fruits [Chillet et al., 2013]).

Les analyses de sensibilité du module plante permettent d'évaluer l'influence de caractères génétiques et agronomiques de la plante dans un contexte d'infestation moyen. Sachant que les nécroses présentes sur la plante provoquent des dégâts considérables à la récolte [Castelan et al., 2013], nous avons choisi de présenter une seule sortie d'intérêt : le pourcentage de nécroses sur la plante aux dates de floraison et 30 jours avant récolte.

Enfin, les analyses de sensibilité du modèle complet SiBaToka permettent d'évaluer l'effet de tous les paramètres du modèle. Nous présenterons, comme précédemment, la variabilité du pourcentage de nécroses sur un plant à deux dates du cycle phénologique de la plante : à la floraison et 30 jours avant la récolte.

Module	Sortie d'intérêt	Date d'intérêt
épidémiologique	nombre de lésions en incubation	
	% surface conidienne	initiation florale
	PJFM	floraison
	% nécroses	30 jours avant la récolte
	PJFN	
plante	% maladie	
	% nécroses	floraison
		30 jours avant la récolte
SiBaToka	% nécroses	floraison
		30 jours avant la récolte

TABEAU 3.5 – Récapitulatif des sorties d'intérêt considérées à différentes dates pour chacun des modules et pour les deux méthodes d'analyse de sensibilité : Morris et e-FAST

Gammes de variation des paramètres

Les gammes de variation nécessaires à la réalisation de ces analyses sont tirées de la littérature, pour les 20 paramètres du module plante, et d'expérimentations en laboratoire ou au champ pour les 11 paramètres du module épidémiologique (Tableaux 3.1 et 3.2). Les valeurs minimales et maximales observées dans les différentes expérimentations sont détaillées dans [Landry et al., in prep].

Ces bornes ont été choisies de manière à représenter la grande diversité des cultivars i.e. des bananiers sensibles et des bananiers présentant différents niveaux de résistances partielles à la cercosporiose noire.

3.3 Résultats

3.3.1 Évaluation de l'effet des composantes de résistance de la plante et de la dispersion de l'inoculum secondaire

Par la méthode de Morris

Les analyses de sensibilité du module épidémiologique réalisées par la méthode de Morris pour chacune des deux situations parasitaires (inoculum primaire faible et élevé), sur trois cycles de culture, pour chacune des sorties d'intérêt et à trois temps du cycle de la plante (cf. [Annexes A.1 à A.12](#)) permettent d'une part d'observer des effets différents suivant la dose d'inoculum primaire proposée et d'autre part d'évaluer temporellement l'effet des paramètres.

Effet des composantes de résistance des bananiers. Globalement, pour toutes les sorties d'intérêt, pour tous les temps considérés et pour les deux pressions parasitaires, les paramètres liés aux composantes de résistance les moins influents sont ceux liés à la sporulation asexuée en particulier sc_a (la capacité de sporulation) et a_{min} (la taille minimale d'une lésion émettrice de conidies).

Plus précisément, concernant la première étape du cycle épidémique ; l'étape de création et de croissance des lésions, la variabilité des pourcentages de maladie sur la plante 30 jours avant la récolte en fonction des 6 valeurs du paramètre lgr (vitesse d'extension des lésions) indique que des vitesses de croissance très lentes (resp. rapides) entraînent un très faible (resp. fort) pourcentage de maladie sur la plante pour les trois cycles de culture (Figure 3.2). A vitesse égale, les pourcentages de maladie sur la plante diffèrent selon la pression d'inoculum primaire : il y a moins de maladie lorsque l'inoculum est faible. Ces différences s'observent nettement en premier cycle. Ce paramètre semble donc jouer un rôle majeur dans la quantité de maladie présente sur la plante à 30 jours avant la récolte.

Plus globalement, la vitesse d'extension des lésions se classe dans chacun des cas étudiés, dans les trois premiers paramètres ayant les plus fortes valeurs de μ^* (cf. [Annexes A.1 à A.12](#)).

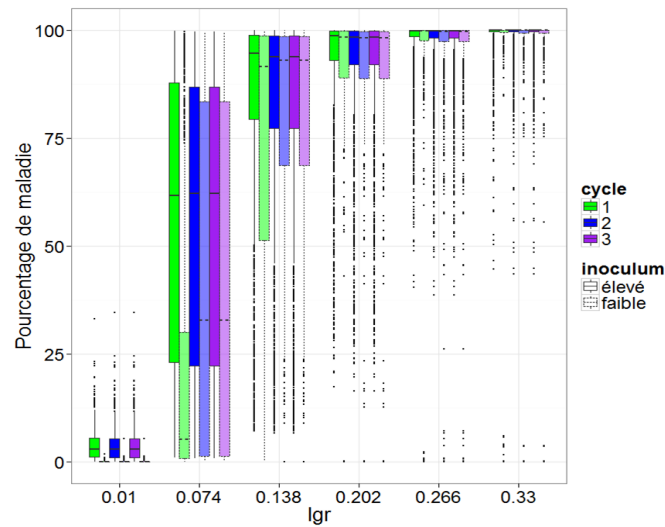


FIGURE 3.2 – Pourcentage de maladie par plante sur les trois premiers cycles de culture sous deux pressions d'inoculum contrastées en fonction des six valeurs du paramètre lgr (vitesse d'extension des lésions).

Dans l'étape de sporulation, l'étude plus détaillée de la variabilité des pourcentages de maladie sur la plante 30 jours avant la récolte en fonction des 6 valeurs du paramètre sc_s (capacité de sporulation sexuée) indique que quelle que soit la quantité d'ascospores produites, les valeurs maximales de maladie sont proches de 100% (Figure 3.2). La variabilité de la sortie semble diminuer cependant à mesure que le nombre de spores produites augmente.

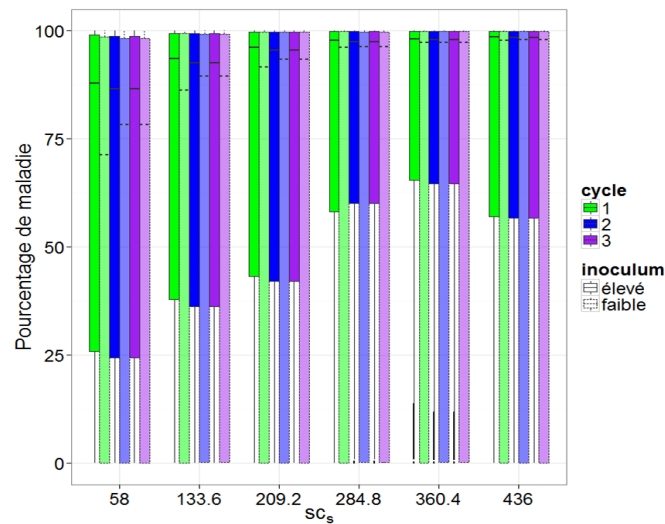


FIGURE 3.3 – Pourcentage de maladie par plante sur les trois premiers cycles de culture avec deux pressions d'inoculum contrastées en fonction des six valeurs du paramètres sc_s (capacité de sporulation sexuée).

Effet de la dispersion de l'inoculum secondaire. Globalement, pour toutes les sorties d'intérêt, tous les temps considérés et pour les deux pressions parasitaires, le paramètre β_s lié à la dispersion sexuée est systématiquement considéré comme négligeable. Les paramètres β_a lié à la dispersion asexuée et sc_a , décrivant la capacité de sporulation asexuée (production de conidies) sont aussi classés parmi les paramètres les moins influents i.e. ceux ayant de faibles valeurs de μ^* . Dans

l'étape de dispersion, nous nous intéressons à la variance des pourcentages de maladie sur la plante 30 jours avant la récolte, comme pour les capacités de sporulation, en fonction cette fois des 6 valeurs du paramètre β_s (paramètre de la fonction de dispersion sexuée). La variabilité de la sortie ne semble pas être corrélée à la valeur du paramètre sur son domaine de variation. En effet, quelle que soit la valeur de β_s , la maladie peut atteindre 100% et peut avoir des valeurs minimales très contrastées (Figure 3.4).

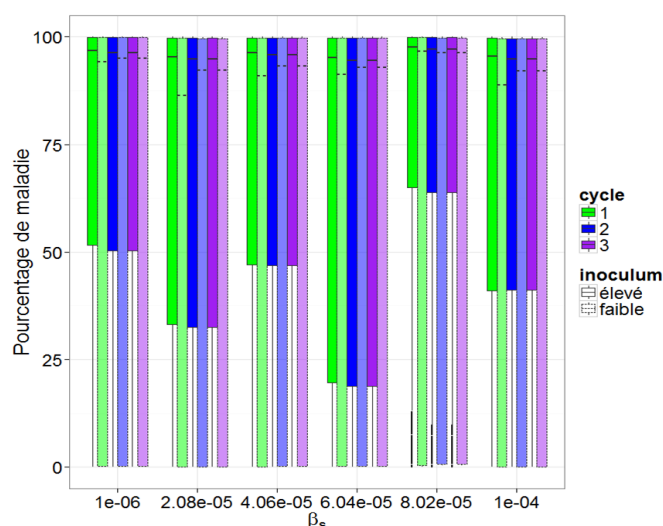


FIGURE 3.4 – Pourcentage de nécroses sur les trois premiers cycles de culture avec deux pressions d'inoculum contrastées en fonction des six valeurs du paramètre β_s (paramètre de la fonction de dispersion des spores sexuées)

Ce module ayant peu de paramètres (11) nous avons choisi de tous les sélectionner pour l'étape suivante d'analyse : la méthode e-FAST.

Par la méthode e-FAST

Les effets totaux et principaux des 11 paramètres calculés pour chacune des sorties d'intérêt aux trois dates du développement phénologique de la plante et pour les deux situations d'inoculum primaire permettent d'identifier les trois paramètres les plus influents : lgr (vitesse d'extension des lésions), ei (efficacité d'infection d'une feuille nouvellement émise) et di (durée d'incubation) (Figure 3.5 pour la sortie d'intérêt pourcentage de maladie sur la plante et cf. Annexes A.17 à A.21 pour les autres sorties). Ils ont tous les trois liés à l'étape de création et de croissance des lésions du cycle infectieux.

Compte tenu de la grande influence de ces trois paramètres, nous proposons leurs analyses sur cinq dates du cycle phénologique de la plante (Figures 3.6, 3.7 et 3.8).

Le paramètre lgr ressort comme le paramètre le plus influent quel que soit le stade de développement de la plante, la situation parasitaire et le cycle de culture. Ses variations indépendamment de la valeur des autres paramètres (effet principal) expliquent environ 70% de la variance du pourcentage de surface malade 30 jours avant la récolte quel que soit l'inoculum présent dans l'air. Avec

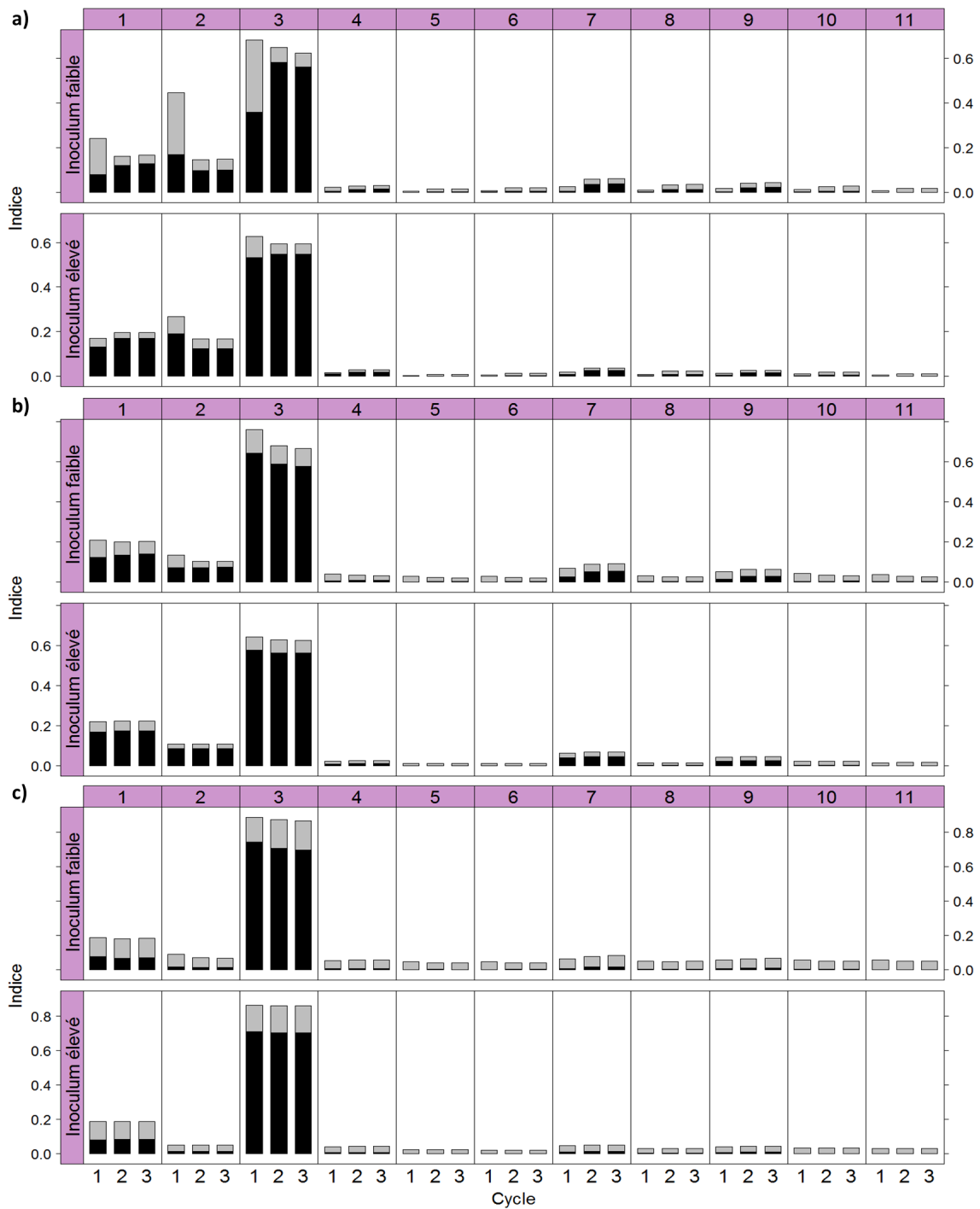


FIGURE 3.5 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 11 paramètres du module épidémiologique calculés sur le pourcentage de maladie sur la plante à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Numérotation des paramètres : 1: ei , 2: ip , 3: lgr , 4: ei^{age} , 5: a^{min} , 6: sc_a , 7: sc_s , 8: Δ_a , 9: Δ_s , 10: β_a , 11: β_s .

ses interactions avec les autres paramètres (effet total) ses variations expliquent quasiment 90% de la variation du pourcentage de surface malade à cette même date. L'évaluation à différentes dates des effets de ce paramètre (Figure 3.6) suggère que son influence croît dans le temps (avec ou sans interactions) et reste identique en deuxième et troisième cycle.

Le paramètre ip ressort comme un des plus influents principalement en début de cycle (Figure 3.7). On observe une nette décroissance de son effet au cours du temps. À l'initiation florale, pour les trois cycles et les deux pressions parasitaires, les variations du paramètre ip expliquent environ 15 % de la variance du pourcentage de maladie sur la plante. En fin de cycle, ce pourcentage est quasi-nul.

Le paramètre ei semble être influent tout au long du cycle phénologique de la plante, en cycle deuxième et troisième cycle (Figure 3.8). On observe une nette décroissance de son effet au cours du temps en premier cycle lorsque l'inoculum est faible. À la floraison, pour les trois cycles et deux pressions parasitaires, les variations du paramètre ei expliquent environ 22 % de la variance du pourcentage de maladie sur la plante. C'est à cette date que son influence est la plus élevée.

Hormis les trois paramètres présentés, en début de cycle épidémique, trois autres paramètres présentent une valeur $ST_i > 0.1$ et ce jusqu'à la floraison. Il s'agit des paramètres ei^{age} et des paramètres liés à la sporulation sexuée sc_s et Δ_s (resp. décroissance de la réceptivité de la feuille à l'infection en fonction de son âge, capacité de sporulation sexuée et durée de sporulation sexuée).

En milieu de cycle épidémique, la variabilité des surfaces conidiennes simulées est très influencée par le paramètre Δ_a (durée de sporulation conidienne) qui explique à la floraison jusqu'à 50% de la variance de cette sortie. Le paramètre sc_a (capacité de sporulation asexuée) explique aussi environ 10% la variance des surfaces conidiennes à cette même date et 30 jours avant la récolte. Tous les paramètres influent sur la variance du rang de la PJFM par le biais de fortes interactions. Hormis les trois paramètres décrits, seul le paramètre a_{min} (surface minimale d'une lésion produisant des conidies) explique la variance des PJFM indépendamment de la valeur des autres paramètres (effet principal) mais cela dans une proportion très faible ($S_i < 0.1$).

En fin de cycle épidémique, parmi les autres paramètres, seul Δ_a explique environ 10% de la variance du pourcentage de nécroses sur la plante. Comme précédemment, tous les paramètres influent sur la variance du rang de la PJFN par le biais de fortes interactions, seul le paramètre Δ_a (durée de sporulation asexuée) explique environ 20% de la variance du rang sans interaction avec les autres paramètres.

Les paramètres β_s , β_a , Δ_s peuvent être considérés comme non influents tout au long du cycle épidémique et du cycle phénologique de la plante. Leurs effets principaux sont systématiquement inférieurs à 0.1, de même pour leurs effets totaux $ST_i < 0.1$ hormis pour les sorties PJFM et PJFN.

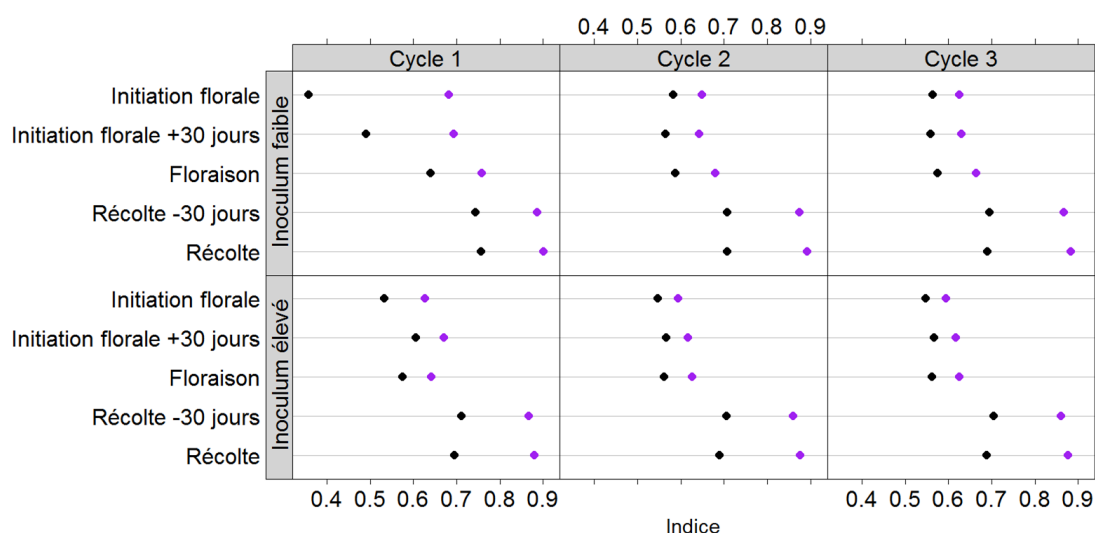


FIGURE 3.6 – Indices de sensibilité principaux (en noir) et totaux (en violet) du paramètres 3 : lgr du module épidémiologique calculés sur le pourcentage de maladie sur la plante à cinq temps selon deux doses d'inoculum.

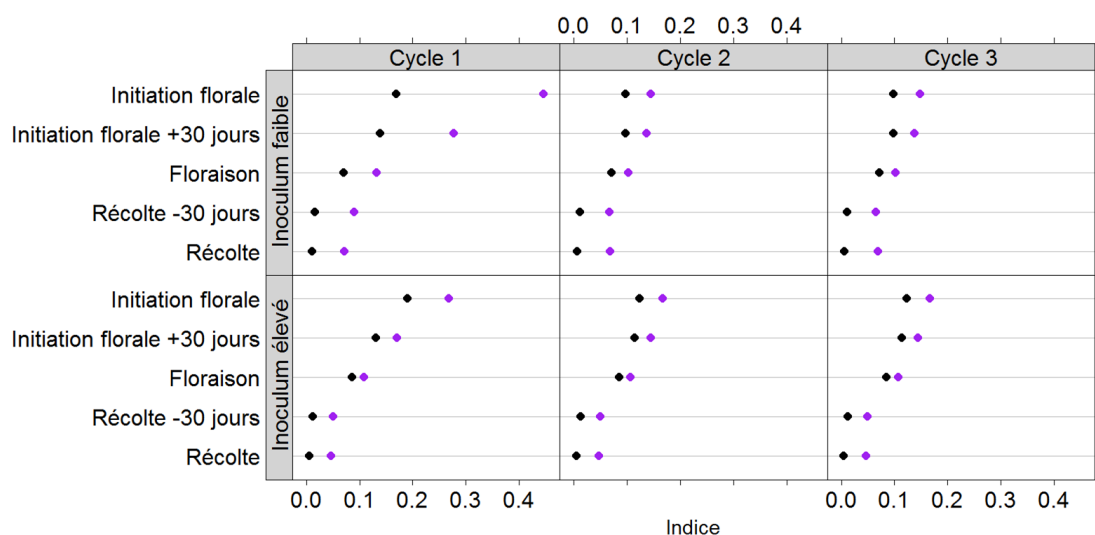


FIGURE 3.7 – Indices de sensibilité principaux (en noir) et totaux (en violet) du paramètres 2 : ip du module épidémiologique calculés sur le pourcentage de maladie sur la plante à cinq temps selon deux doses d'inoculum.

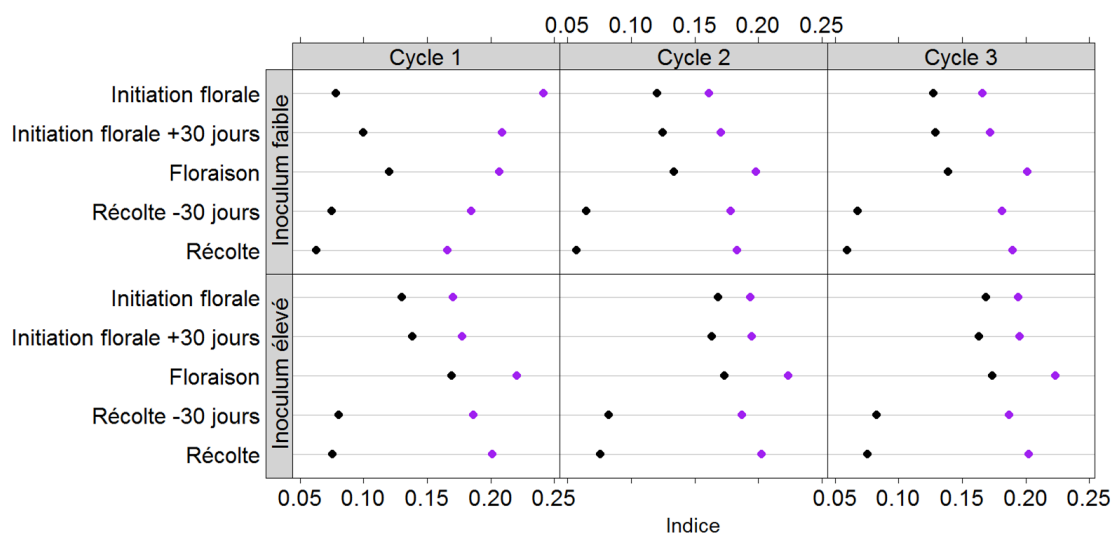


FIGURE 3.8 – Indices de sensibilité principaux (en noir) et totaux (en violet) du paramètres 1 : ei du module épidémiologique calculés sur le pourcentage de maladie sur la plante à cinq temps selon deux doses d'inoculum.

3.3.2 Évaluation de l'effet des caractères agronomiques et génétiques de la plante

Par la méthode de Morris

L'analyse de sensibilité du module plante par la méthode de Morris est présentée seulement pour une seule sortie d'intérêt (le pourcentage de surface nécrosées par plante) et à deux dates du cycle de la plante (floraison et 30 jours avant la récolte) (Tableau 3.5).

Cette analyse permet d'une part d'observer des effets différents suivant la dose d'inoculum primaire fixée et d'autre part d'évaluer à différentes dates l'effet des paramètres. Toutes les figures sont présentées en [Annexes A.13 et A.14](#).

Globalement, parmi tous les caractères agronomiques et génétiques qui caractérisent la plante, les paramètres décrivant les hauteurs initiales d'une plante ou de son rejet (H_L^1 et H_L^0) sont considérés comme négligeables (ces paramètres ayant de faibles valeurs de μ^*). Certains paramètres ont une influence contrastée dans le temps, par exemple le nombre de feuilles sur la plante à la récolte (paramètre $N_L^{harvest}$) a peu d'influence jusqu'à la floraison (rang 19 sur 20 selon sa valeur moyenne de μ^*) et devient parmi ceux les plus influents (rang 5 sur 20) après cette date. On observe également très peu de différences d'influence des paramètres entre le deuxième et le troisième cycle de culture (cf. [Annexes A.13 et A.1](#)).

L'étude détaillée de la distribution des pourcentages de nécroses sur la plante 30 jours avant la récolte en fonction des valeurs prises par les paramètres génétiques et agronomiques de la plante met en évidence les deux effets suivants :

- le pourcentage de nécroses augmente avec N_L^{flower} , nombre de feuilles émises avant la floraison (Figure 3.9), paramètre génétique classé au quatrième rang ; une floraison précoce entraîne donc un faible pourcentage de nécroses.
- le pourcentage de nécroses diminue avec ler^{final} , rythme d'émission foliaire d'une plante adulte (Figure 3.10), paramètre agronomique classé au deuxième rang ; un rythme rapide entraîne donc un faible pourcentage de nécroses.

Ces deux effets s'observent sur les trois cycles et sont plus marqués en faible pression parasitaire.

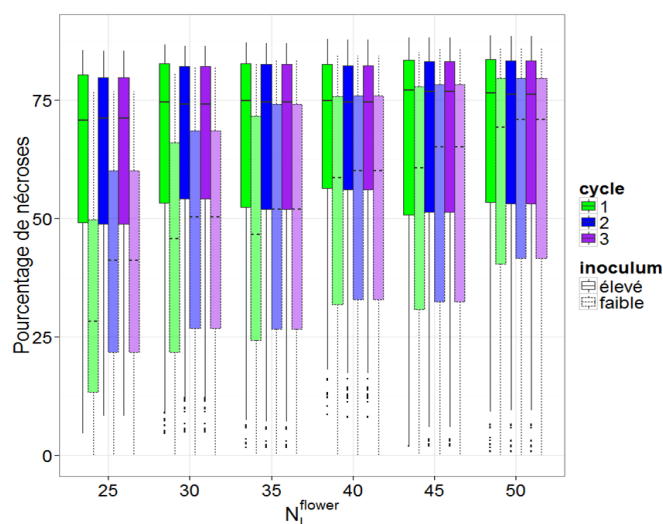


FIGURE 3.9 – Pourcentage de nécroses par plante sur les trois cycles de cultures sous deux pressions d'inoculum contrastées en fonction des six valeurs du paramètre N_L^{flower} (nombre de feuilles émises avant la floraison).

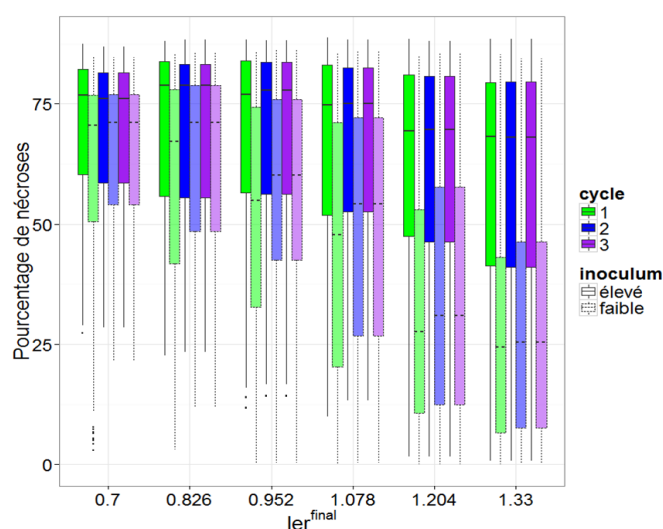


FIGURE 3.10 – Pourcentage de nécroses par plante sur les trois cycles de cultures sous deux pressions d'inoculum contrastées en fonction des six valeurs du paramètre ler^{final} (rythme d'émission foliaire d'une plante adulte).

Nous avons choisi de garder dans l'analyse e-FAST les dix paramètres présentant globalement des valeurs de μ^* les plus élevées. Parmi ces dix paramètres, cinq décrivent des caractères génétiques et les cinq autres décrivent des caractères agronomiques (Tableau 3.1).

Par la méthode e-FAST

Compte tenu des temps de calculs et des résultats préliminaires obtenus avec la méthode de Morris mettant en évidence que l'influence des paramètres varie peu entre le deuxième et troisième cycle, nous avons poursuivi l'analyse suivant la méthode e-FAST en considérant seulement deux cycles de culture.

Les effets totaux et principaux des dix paramètres sélectionnés ont été calculés pour le pourcentage de nécroses sur la plante à deux dates et pour les deux situations d'inoculum primaire (Tableau 3.5). Ces indices permettent d'identifier six paramètres les plus influents (ceux ayant des indices totaux > 0.1) : N_L^0 (nombre de feuilles émises à l'apparition du rejet sélectionné), N_L^{flower} (nombre de feuilles à la floraison), ler^{final} (rythme d'émission foliaire final), N_L^{max} (le nombre de feuilles maximal sur la plante), $N_L^{1,init}$ le nombre de feuilles du rejet à la récolte du pied mère et Δ_{fh} le temps entre la floraison et la récolte (Figure 3.11).

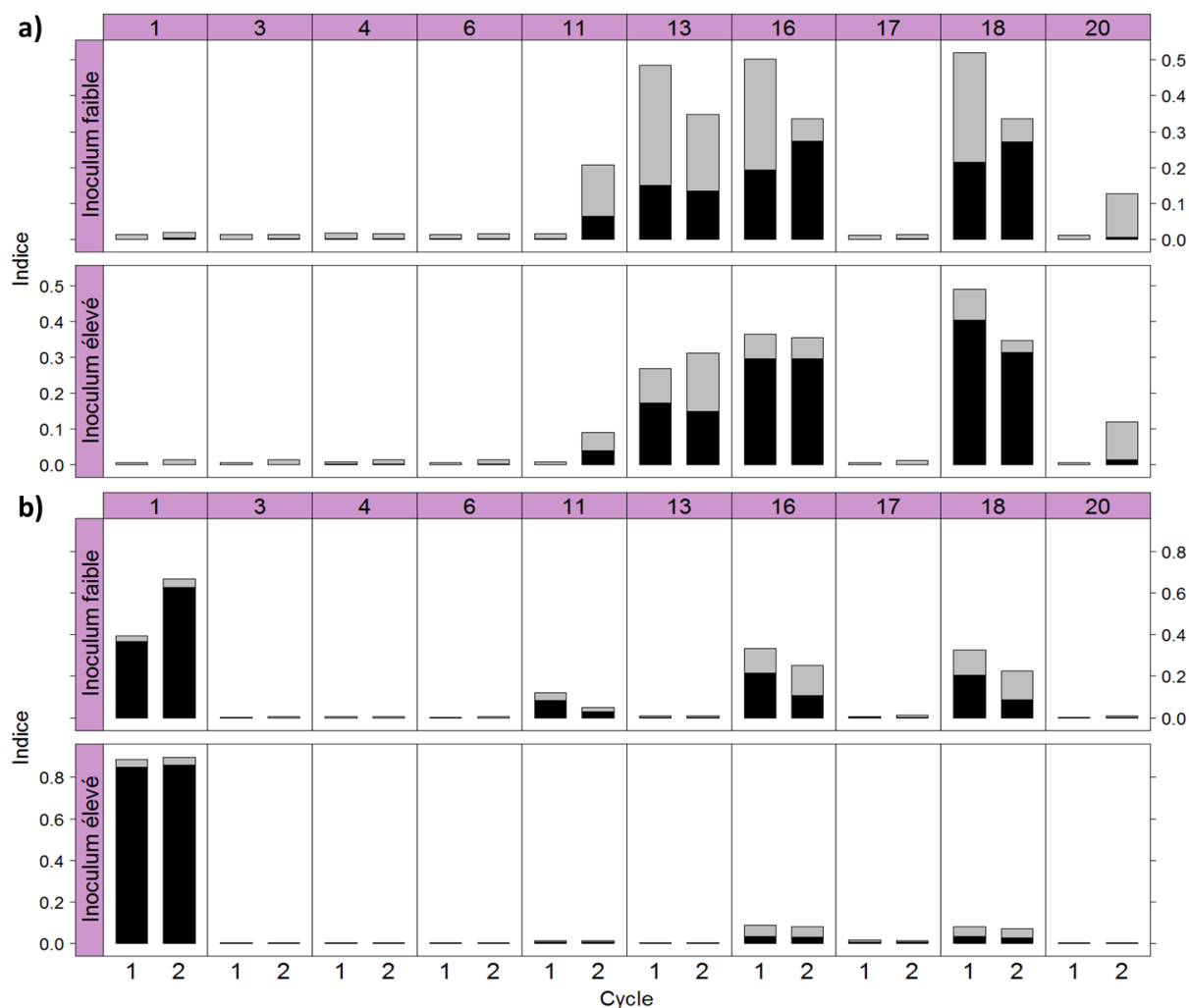


FIGURE 3.11 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 10 paramètres du module plante sélectionnés après l'analyse de Morris calculés sur le pourcentage de nécroses sur la plante à deux dates, a) floraison et b) 30 jours avant la récolte. Numérotation des paramètres : 1 : Δ_{fh} , 3 : N_L^{max} , 4 : N_L^{min} , 6 : W_L^{min} , 11 : N_L^{flower} , 13 : N_L^0 , 16 : ler^{final} , 17 : $N_L^{harvest}$, 18 : N_L^{max} , 20 : $N_L^{1,init}$

Parmi ces six paramètres influents, deux sont classés comme caractères génétiques (1 : Δ_{fh} et 11 : N_L^{flower}) et quatre comme caractères agronomiques (13 : N_L^0 , 16 : ler^{final} , 18 : N_L^{max} , 20 : $N_L^{1,init}$) (Tableau 3.1).

Aux deux temps considérés, deux paramètres influent sur la variance du pourcentage de nécroses : les paramètres agronomiques ler^{final} et N_L^{max} . Lorsque l'inoculum primaire est faible, les

variations de ces paramètres expliquent jusqu'à 50% de la variance du pourcentage de nécroses à la floraison du premier cycle et 30% de cette variance 30 jours avant la récolte. Ces taux baissent en deuxième cycle et lorsque l'inoculum est élevé.

À la floraison, le paramètre agronomique N_L^0 peut être considéré comme très influent. Ses variations entraînent une variance de 50% du pourcentage de nécroses en premier cycle lorsque l'inoculum est faible et environ 25% lorsque l'inoculum est élevé. En deuxième cycle, les tendances s'inversent, l'influence du paramètre baisse lorsque l'inoculum est faible et augmente lorsque l'inoculum est élevé.

Deux autres paramètres influent sur la variance du pourcentage de nécroses à la floraison : le paramètre génétique N_L^{flower} et le paramètre agronomique $N_L^{1,init}$. Il est à préciser que ces deux paramètres sont influents seulement en deuxième cycle et le sont d'autant plus que l'inoculum primaire est faible.

Le paramètre génétique Δ_{fh} , comme le montrent les résultats de l'analyse selon la méthode de Morris, est très influent 30 jours avant la récolte : lorsqu'en deuxième cycle l'inoculum est faible, sa variabilité explique jusqu'à 60% de la variance de la sortie d'intérêt et dépasse les 80% pour les deux cycles lorsque l'inoculum est élevé.

3.3.3 Évaluation du modèle complet

Par la méthode de Morris

Les résultats de la méthode de Morris sur les sorties d'intérêt décrites pour le modèle complet SiBaToka (Tableau 3.5) obtenus pour les trois cycles de culture et les deux pressions d'inoculum permettent de conclure que globalement, les paramètres les moins influents sont ceux caractérisant la croissance de la plante. Les paramètres du module épidémiologique sont, quant à eux, classés dans les 11 premiers paramètres d'après leurs valeurs de μ^* (cf. Annexes A.15 et A.16). On remarque aussi que l'influence des paramètres est identique en deuxième et troisième cycle.

Les résultats sont similaires à ceux obtenus dans l'analyse du module plante : les deux paramètres ayant des valeurs de μ^* les plus faibles sont les paramètres décrivant les hauteurs initiales d'une plante ou de son rejet (H_L^0 et H_L^0).

Comme décrit précédemment, certains paramètres ont une influence contrastée dans le temps, ici la durée entre la floraison et la récolte (paramètre Δ_{fh}) a peu d'influence jusqu'à la floraison (rang 30/31 selon sa valeur moyenne de μ^*) et devient parmi ceux les plus influents (rang 3/31) après cette date (cf. Annexes A.15 et A.16).

L'étude détaillée la variabilité des pourcentages de nécroses sur la plante 30 jours avant la récolte, en fonction des six valeurs du paramètre Δ_{fh} (durée entre la floraison et la récolte), classé au troisième rang, montre que le pourcentage de nécroses est d'autant plus élevé que cette durée est longue. Comme précédemment, ces différences s'observent plus nettement en faible pression parasitaire (Figure 3.12).

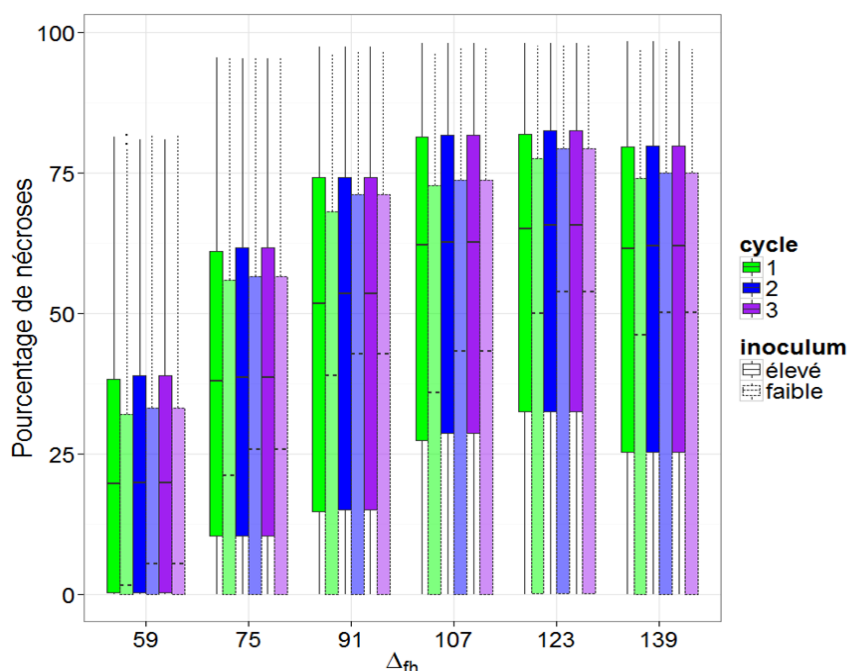


FIGURE 3.12 – Pourcentage de nécroses sur les trois cycles de cultures avec deux pressions d'inoculum contrastées en fonction des six valeurs du paramètres Δ^{fh} (temps entre la floraison et la récolte).

Nous avons choisi de garder dans l'analyse e-FAST les 15 paramètres présentant globalement des valeurs de μ^* les plus élevées. Ces paramètres comprennent dix paramètres du module épidémiologique et cinq paramètres du module plante (Tableau 3.1).

Par la méthode e-FAST

Comme précédemment, compte tenu des temps de calcul et des résultats préliminaires obtenus avec la méthode de Morris mettant en évidence que l'influence des paramètre varie peu entre le deuxième et troisième cycle, nous avons poursuivi l'analyse suivant la méthode e-FAST en considérant deux cycles de culture.

Les effets totaux et principaux des 15 paramètres sélectionnés ont été calculés pour le pourcentage de nécroses sur la plante à deux dates d'un cycle de culture et pour les deux situations d'inoculum primaire (Tableau 3.5).

Ces indices permettent d'identifier six paramètres les plus influents (ceux ayant des indices totaux > 0.1), trois paramètres du module épidémiologique : ei (efficacité d'infection d'une feuille nouvellement émise), ip (période d'incubation), lgr (taux de croissance des lésions), et trois paramètres du module plante : Δ_{fh} (temps entre la floraison et la récolte), ler^{final} (rythme d'émission foliaire final) et N_L^{max} (nombre de feuilles maximal sur la plante) (Figure 3.13). Parmi les trois paramètres du module plante, le premier est à caractère génétique, les deux suivants sont à caractère agronomique.

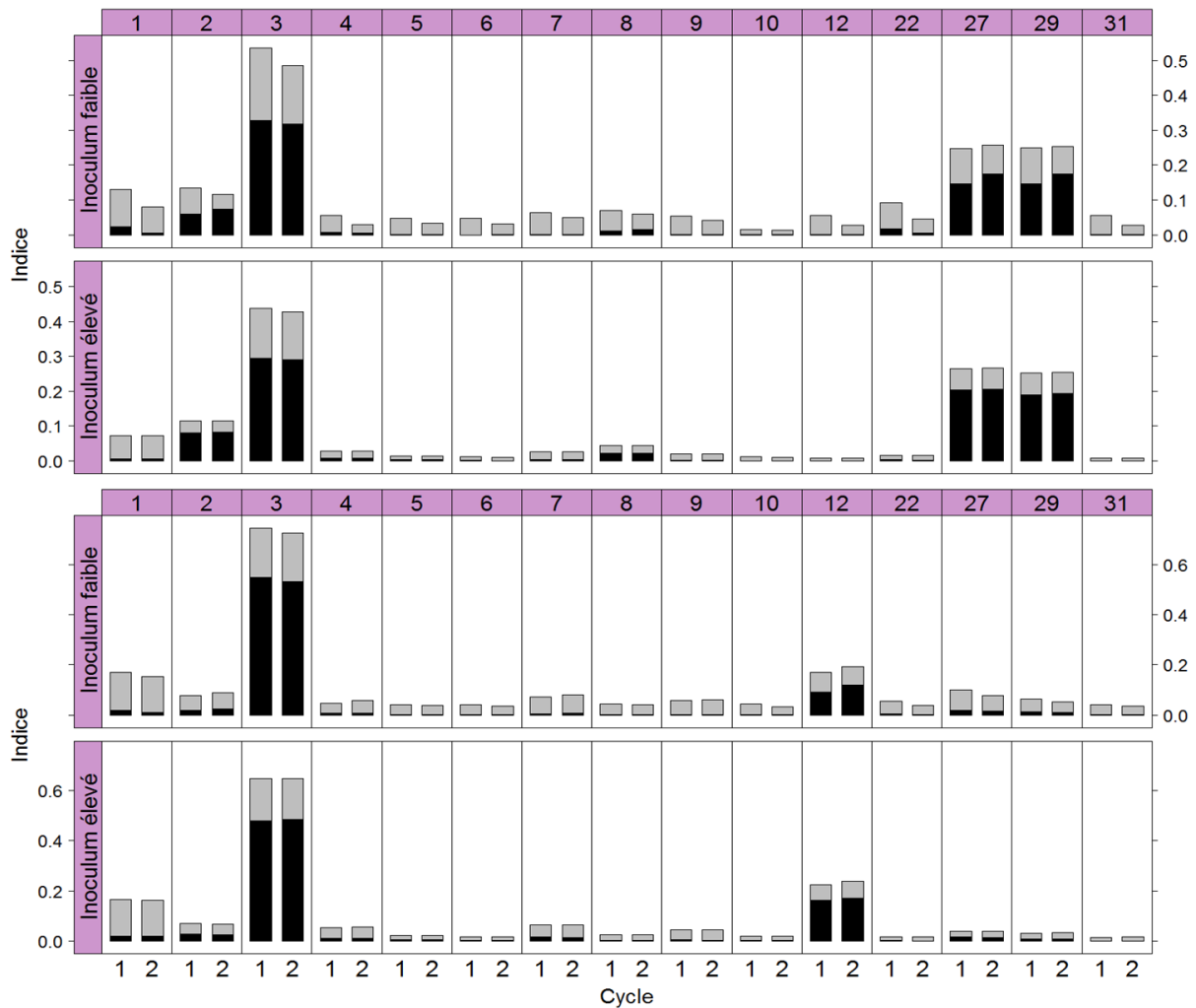


FIGURE 3.13 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 15 paramètres du modèle complet SiBaToka sélectionnés après l'analyse de Morris calculés sur le pourcentage de nécroses sur la plante à deux dates, a) floraison et b) 30 jours avant la récolte. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 12 : Δ_{fh} , 22 : N_L^{flower} , 27 : ler^{final} , 29 : N_L^{max} , 31 : $N_L^{1,init}$

Les résultats de cette analyse confirment que les variabilités des trois paramètres du module épidémiologique influent sur la variance du pourcentage de nécroses à la floraison et 30 jours avant la récolte. La variabilité du paramètre lgr explique jusqu'à 50% de la variance du pourcentage de nécroses à la floraison (et plus de 40% en inoculum faible) et plus de 60% 30 jours avant la récolte quelle que soit la pression parasitaire. En inoculum élevé, l'effet des deux paramètres ip et ei semble équivalent pour les deux cycles. En inoculum faible, l'influence de ces paramètres diminue. On remarque aussi que le paramètre ip a une valeur d'indice principal (effet du paramètre seul) proche de 0.1 à la floraison et que cet indice diminue fortement 30 jours avant la récolte. Le paramètre ei a une valeur d'indice total négligeable ($ST_i < 0.1$) à la floraison et proche de 0.2 30 jours avant la récolte. L'influence de ce paramètre est cependant pour les deux dates principalement due aux interactions.

On retrouve dans l'analyse du modèle complet les paramètres du module plante considérés comme les plus influents : la variabilité de ler^{final} et N_L^{max} à la floraison explique environ 25% de la variance du pourcentage de nécroses et la variabilité de Δ_{fh} en explique environ 20% 30 jours avant la récolte.

3.4 Discussion

En raison de la complexité du modèle SiBaToka et de sa structure en deux modules, nous avons réalisé deux types d'analyses de sensibilité sur chacun de ses sous modules (épidémiologique et plante) et sur le modèle complet. Dans une première approche, nous avons utilisé des analyses quantitatives suivant la méthode de Morris nous permettant d'identifier les paramètres les moins influents dans chacun des modules étudiés afin de les supprimer dans les analyses quantitatives e-FAST effectuées en deuxième approche. Nous revenons en détail ici sur les trois modules analysés.

3.4.1 Discussion des principaux résultats

Trois composantes de résistance de la plante aux influences majeures

Les analyses du module épidémiologique ont montré la forte influence, toutes sorties confondues, de paramètres épidémiologiques qui correspondent à trois composantes de résistance : la vitesse d'extension des lésions, la durée d'incubation et l'efficacité d'infection.

L'influence de la durée d'incubation sur d'autres pathosystèmes est établie. Par exemple, [Leclerc et al. \[2014\]](#) montrent la grande influence de la durée d'incubation dans le processus épidémiologie de *Rhizoctonia solani* sur la betterave. L'influence de la durée d'incubation mise en évidence sur la sévérité de la cercosporiose noire est en accord avec les résultats expérimentaux décrit par [Fouré et al. \[1984\]](#) sur plusieurs cultivars. Les auteurs signalent cependant un manque de fiabilité de ce paramètre pour déterminer la sensibilité d'une variété. Nous montrons, en effet, une décroissance de l'influence de ce paramètre au cours d'un cycle phénologique de la plante. Sous l'hypothèse d'une décroissance de la réceptivité d'une feuille à l'infection avec son âge, le nombre de nouvelles infections diminue aussi et par conséquent la durée d'incubation du champignon influe peu en fin de cycle.

Concernant la vitesse d'extension des lésions et l'efficacité d'infection qui nécessitent des mesures assez lourdes, nous n'avons pas de connaissances précises sur le poids épidémiologique de ces paramètres.

L'influence de la vitesse d'extension des lésions sur la sévérité de la cercosporiose noire a été suggérée par des approches expérimentales (par comparaison du niveau de maladie entre une variété sensible et des variétés hybrides résistantes dont certaines composantes de résistance avaient été mesurées) [[Abadie et al., 2009](#)]. De plus, [Fouré et al. \[1984\]](#) ont mesuré la durée de développement des lésions sur différentes variétés et considèrent ce paramètre comme un critère fiable pour

déterminer la sensibilité d'une variété. Sur d'autres pathosystèmes, la vitesse d'extension des lésions a aussi été mise en évidence par [Caubel \[2012\]](#) dans l'analyse de sensibilité du modèle générique développé (MDMA) couplé au modèle de culture STICS calibré pour le mildiou et l'oïdium de la vigne.

L'influence de l'efficacité d'infection sur le niveau de maladie n'était pas clairement démontré dans le cas de notre pathosystème, mais des observations épidémiologiques suggéraient son influence dans le cycle épidémique de la maladie [[Abadie et al., 2003](#)]. Nous avons pu ici évaluer l'importance de ce paramètre en considérant une croissance optimale de la plante (sur le module épidémiologique) mais aussi en interaction avec tous les paramètres du modèle.

Inoculum primaire et inoculum secondaire

Comme le montrent [Hong and Hwang \[1998\]](#) qui décrivent une relation linéaire entre la dose d'inoculum primaire déposée et la quantité de maladie résultante de l'infection de *Colletotrichum coccodes* sur des plants de poivre, on observe sur notre pathosystème des quantités de maladie plus importantes lorsque l'inoculum primaire est élevé. L'influence des paramètres sur les sorties d'intérêt va elle aussi être différente selon l'inoculum primaire considéré (quantités de maladie plus faibles), mais les paramètres les plus influents le sont quelle que soit la pression parasitaire. Les différences s'observent d'ailleurs plus nettement lors du premier cycle de culture, et l'influence des paramètres sur les variances des sorties d'intérêt semble être la même pour les cycles suivants. [Mobambo et al. \[1996\]](#) montrent, en effet, que le niveau de maladie sur plantain est plus faible en premier cycle qu'en second cycle.

Bien que les phytopathogènes polycycliques aient tendance à avoir de forts taux d'auto-infection i.e. de forts taux de production d'inoculum secondaire se déposant sur la même plante ([Lannou et al. \[2008\]](#), [Frezal et al. \[2009\]](#)), la production de spores (capacités de sporulation sexuée et asexuée) semble, dans l'étude de ce modèle, peu influente sur la quantité de maladie (et de nécroses) présente sur la plante sachant les gammes de variations de paramètres établies. Cependant, la variance du nombre de lésions en incubation semble être influencée par la production sexuée des nécroses, certainement de par leurs fortes capacités de sporulation et la fonction de dispersion isotrope du modèle. Ce résultat renseigne sur la composition de l'inoculum secondaire. En effet, ce sont les paramètres de la sporulation sexuée (sc_s , capacité et Δ_s , durée) et non ceux de la sporulation asexuée qui influent sur le nombre de lésions en incubation. Ces résultats suggèrent que l'inoculum secondaire serait essentiellement composé d'ascospores. En effet, dans nos cas d'études, la variance du nombre de spores asexuées produites ne semble pas influencer les sorties d'intérêt considérées.

Caractéristiques agronomiques et génétiques de la plante

Les principaux paramètres à caractère agronomique influents sur le pourcentage de nécroses sur la plante suggèrent que maintenir un rythme d'émission foliaire rapide (par le biais de bonnes pratiques culturales) et pratiquer un effeuillage sanitaire des feuilles nécrosées (réduisant leurs nombre sur la plante) sont importants pour réduire la quantité de nécroses, surtout en premier cycle et lorsque l'inoculum primaire est faible. Ces résultats sont en accord avec [Mobambo et al., 1994] qui montrent l'effet du rythme d'émission foliaire sur le pourcentage de nécroses de la cercosporiose noire d'une variété plantain par le biais d'une enquête, et Marin et al. [2003] qui recommandent la réduction de la quantité d'inoculum secondaire de la plante, par le biais d'effeuillages, pour réduire efficacement la réinfestation.

Les paramètres génétiques les plus influents à la date établie comme importante pour l'impact de la maladie sur la production de fruits [Chillet et al., 2013] sont les paramètres relatifs aux durées de cycle. La durée entre la floraison et la récolte, période pendant laquelle la plante n'émet plus de feuilles, est le paramètre le plus influent dans l'analyse du module plante et dans l'analyse du modèle complet. Or, ce résultat peut s'expliquer par une durée d'infestation plus longue dans le cas des cycles plus longs. Cependant, ce résultat peut être considéré comme une information nouvelle car la littérature ne mentionne aucune étude permettant de relier la durée entre la floraison et la récolte et le niveau de maladie. Le deuxième paramètre influent est le nombre de feuilles émises avant la floraison, soit le nombre total de feuilles émises par la plante. En effet, plus le bananier produit de feuilles, plus la durée du cycle est longue et plus le pourcentage de nécroses est élevé. Même si, à notre connaissance, aucune étude ne s'intéresse à la durée des cycles de la plante sur la sévérité de la cercosporiose noire, Noupadja et al. [2007] mentionne pour des variétés différentes des niveaux de maladie similaires pour des intervalles entre la plantation et la floraison.

3.4.2 Discussion des méthodes et du modèle

Méthodes utilisées et perspectives

L'utilisation d'une méthode d'analyse de sensibilité qualitative suivie d'une méthode quantitative a permis de minimiser les temps de calcul en supprimant les paramètres peu influents détectés par la première méthode et de calculer les indices de sensibilité principaux et totaux, interprétables en terme de pourcentages de variance expliquée, pour les paramètres considérés comme les plus influents.

Dans le cas de l'analyse du module épidémiologique, les résultats des deux méthodes sont similaires : les paramètres classés les plus influents selon leurs valeurs de μ^* sont aussi les paramètres ayant des indices totaux les plus élevés, ceci appuyant la validité de nos résultats.

La méthode e-FAST a permis d'évaluer les indices principaux S_i et les indices totaux ST_i de chaque paramètre i . Cependant, concernant les composantes de résistance par exemple, van den Berg et al. [2013] montrent qu'avec une forte capacité d'auto-infection de la plante, due à une

forte densité de lésions, la période de latence de la rouille brune du blé est raccourcie. Sur notre pathosystème, [Fouré et al. \[1984\]](#) et [Mobambo et al. \[1996\]](#) suggèrent par exemple, que la densité de lésions présentes sur la plante influe sur la vitesse d'extension des lésions. Il semblerait donc intéressant d'évaluer plus en détail l'interaction entre les paramètres du modèle SiBaToka.

La méthode e-FAST utilisée calcule des effets principaux et totaux, et ne permet pas d'explorer les interactions d'ordre supérieur. L'utilisation d'une méthode de Sobol par exemple (plus coûteuse en temps de calcul) permettrait de calculer précisément des indices d'interaction d'ordre supérieur ([Iooss \[2010\]](#), [Saltelli et al. \[2006\]](#)).

Une autre méthode serait de considérer un plan factoriel complet pour calculer à l'aide d'ANOVA des indices d'interactions d'ordre 2 sur les deux modules comme le décrivent [Lô Pelzer et al. \[2010a\]](#) sur un modèle de culture structuré en plusieurs sous-modèles. Cette méthode coûteuse en temps de calcul mais robuste et facilement réalisable [[Monod et al., 2006](#)] pourrait convenir à notre cas d'étude, car les résultats obtenus sur chacun des modules sont en adéquation avec ceux obtenus sur le modèle complet. Une méthode moins coûteuse en temps de calcul que la précédente serait d'utiliser un plan factoriel fractionnaire (de résolution V), au lieu d'un plan factoriel complet, pour calculer ces mêmes indices sur le modèle complet [[Fabre et al., 2009](#)].

[Lamboni et al. \[2009\]](#) proposent aussi une méthodologie dans le cadre de modèle de culture dynamique (modèle simulant la variation journalière de matière sèche d'une culture de blé sur 200 jours environ) avec plusieurs sorties d'intérêt qu'il serait intéressant de mettre en place. Le calcul des indices de premier ordre et d'ordre 2 est réalisé à chaque pas de temps par analyse de variance sur les résultats d'un plan factoriel. Une ACP est ensuite réalisée (méthode utilisée aussi dans [Lurette et al. \[2009\]](#)) sur les sorties du modèle et les indices S_i et ST_i sont calculés sur les premières composantes principales.

Sensibilité des sorties d'intérêt à la variation des paramètres : conséquences pour l'amélioration du modèle

Dans le modèle de simulation SiBaToka, les paramètres d'entrée sont des paramètres fixés selon les caractéristiques phénologiques et génétiques de la plante, mais aussi selon les valeurs des composantes de résistance et de dispersion.

Les paramètres les plus influents des modules plante et épidémique sont aussi les paramètres les plus influents dans l'analyse du modèle complet SiBaToka. En effet, ce modèle ne tient pas compte d'interactions entre la croissance de l'hôte et le développement du pathogène, hormis la définition d'un paramètre représentant la perte d'efficacité d'infection liée l'âge de la feuille.

[Leclerc et al. \[2014\]](#) montrent par exemple que la durée d'incubation du pathogène *Rhizoctonia solani* de la betterave varie avec l'âge de l'hôte. Dans le cas de la cercosporiose noire du bananier, des suivis épidémiologiques suggèrent l'effet de la maladie sur la croissance de la plante. En effet, [Mobambo et al. \[1993\]](#) et [Craenen and Ortiz \[1998\]](#) montrent que la maladie retarde la date de floraison sur une variété de type plantain. Il serait alors intéressant de pouvoir intégrer une varia-

bilité temporelle de la valeur des paramètres en fonction des connaissances disponibles, comme le proposent [Cunniffe and Gilligan \[2010\]](#) sur un modèle à compartiments SEIR.

Chapitre 4

Modélisation de la dynamique d'invasion de la cercosporiose noire à l'échelle d'un territoire

Sommaire

4.1 Contextes	117
4.1.1 Contexte biologique	117
Distance de dispersion	118
Effet des facteurs environnementaux sur la dynamique d'invasion	118
4.1.2 Contexte méthodologique	119
4.1.3 Objectifs du chapitre	120
4.2 Données	121
4.2.1 Données de présence-absence à l'échelle d'un territoire	121
4.2.2 Données environnementales	123
Parcellaire déclaré	123
Pluviométrie et vent	124
4.2.3 Reconstitution du paysage	125
4.3 Modèle spatio-temporel	127
4.3.1 Description du modèle développé	127
4.3.2 Simulations du modèle	129
4.4 Inférence des paramètres du modèle	133
4.4.1 Vraisemblance, loi a priori et loi a posteriori	133
4.4.2 Algorithmes et validation	134
Algorithmes	134
4.4.3 Validation	137
4.5 Analyse de sensibilité du modèle	142

4.5.1	Méthode de Morris	143
4.5.2	Méthode extended FAST	144
4.6	Inférence sur données réelles : Article en préparation	148
4.7	Discussion du chapitre	162
4.7.1	Discussion des principaux résultats	162
4.7.2	Discussion méthodologique	163

4.1 Contextes

4.1.1 Contexte biologique

La cercosporiose noire a été détectée pour la première fois dans les îles Fidji en 1963. En deux décennies, elle a envahi les zones bananières tropicales et subtropicales de tous les continents : elle a été détectée sur le continent africain en 1978 au Gabon et sur le continent américain en 1972 au Honduras [Jones, 2000] [Robert et al., 2012]. Dans l'archipel des Caraïbes, elle a été détectée pour la première fois à Cuba fin 1989, la source d'introduction probable étant le Honduras. Depuis les années 1990, elle progresse suivant deux fronts de propagation : un Ouest-Est de Cuba à Porto Rico où elle a été détectée en 2004 et un Sud-Nord depuis Trinidad où elle a été détectée en 2003 [Fortune et al., 2005] jusqu'aux Petites Antilles (Figure 4.1). Aux Antilles françaises, la première détection date de 2010 en Martinique [Ioos et al., 2011] et de 2012 en Guadeloupe [DAAF, communiqué du 24/01/2012]. La détection de la maladie au centre de l'île et l'existence de vent du Sud au cours des mois précédents sont des observations en faveur d'une introduction de la maladie depuis Sainte Lucie, l'île à 50 km au Sud de la Martinique et où la maladie avait été détectée depuis janvier 2010 [Abadie C., communication personnelle].

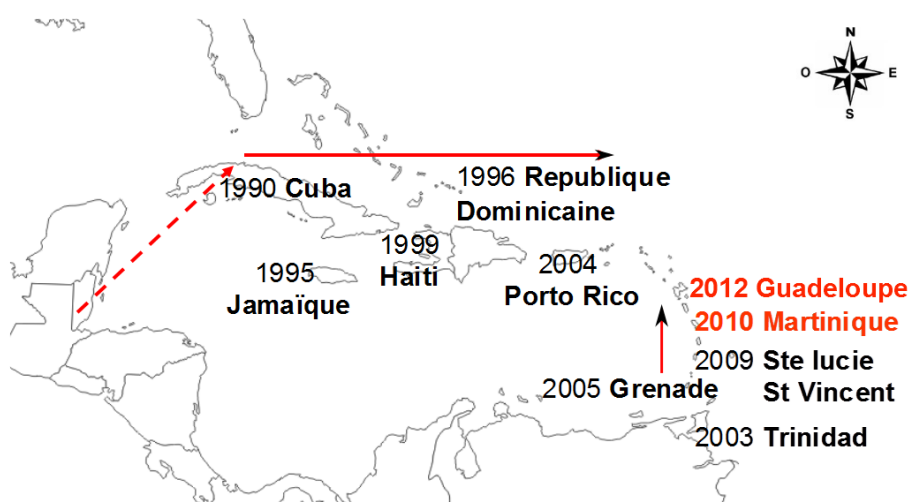


FIGURE 4.1 – Invasion de la cercosporiose noire dans l'arc Caraïbe. Les flèches représentent les fronts de propagation de la maladie.

Comprendre la dynamique d'invasion d'un agent pathogène à l'échelle un territoire est un enjeu contemporain majeur car l'apparition de nouvelles maladies entraîne de graves perturbations des écosystèmes en place, de la biodiversité et de l'économie locale [Fisher et al., 2012]. Une meilleure compréhension passe alors par l'évaluation de la capacité de dispersion à longue distance de l'agent pathogène et des effets des facteurs environnementaux sur cette dispersion [Brown and Hovmøller, 2002].

Distance de dispersion

La distance de dispersion des ascospores du pathogène *Mycosphaerella fijiensis*, spores responsables de la contamination à longue distance, a été étudiée à différentes échelles spatiales et suivant différentes techniques d'expérimentations au champ.

[Rutter et al. \[1998\]](#) proposent d'évaluer la dispersion des ascospores sur une parcelle de 18 m × 18 m (100 cultivars de type plantain espacés de 2 mètres). En positionnant une source infectieuse (feuilles nécrosées) au centre de la parcelle, ils établissent une relation entre l'apparition visuelle des symptômes sur les plants sains et la distance de ces plants à la source infectieuse. Dans cette étude, les auteurs précisent qu'une autre parcelle infectée se trouvait à 100 mètres de leur lieu d'étude et, bien que des pièges à spores placés entre ces deux parcelles n'aient pas capturé de spores, ils n'ont aucune certitude sur le fait que les lésions observées sur la parcelle expérimentale soient provoquées par des spores provenant uniquement de leur source centrale. La même incertitude résulte de l'étude, à plus grande échelle, de [Burt et al. \[1998\]](#) qui concluent, suite à l'apparition de spores dans leurs pièges, à une capture de spores provenant d'une plantation voisine de 40 km. Ils estiment donc la distance de dispersion des ascospores à plus de 40 km.

Pour pallier le manque de connaissance concernant la réelle provenance des spores créant l'infection observée, [Rieux et al. \[2014\]](#) proposent l'estimation d'une fonction de dispersion à l'aide d'une expérience au champ d'environ 2 km × 2 km et d'un suivi génétique des souches de *Mycosphaerella fijiensis* (non présentes dans la zone) utilisées comme source d'infection au centre de la parcelle d'essai. Ils évaluent plusieurs fonctions de dispersion anisotropes et estiment leurs paramètres respectifs. Le meilleur modèle est obtenu avec une fonction de type "exponentielle puissance". La distance moyenne de dispersion (toutes directions confondues) est évaluée à 283 mètres et les distances maximales relevées correspondent à des lésions observées aux bornes de l'expérimentation (1 km de la source). Les ascospores de *Mycosphaerella fijiensis* peuvent donc voyager à plus d'un kilomètre de la source infectieuse comme le montraient déjà [Amil et al. \[2007\]](#) qui détectent la présence du pathogène (identifiable par des analyses moléculaires pour sa résistance à un fongicide spécifique) à plus de 6 km de la zone de production.

Effet des facteurs environnementaux sur la dynamique d'invasion

L'effet des facteurs environnementaux sur les différentes étapes du cycle infectieux, et par conséquent sur le développement des symptômes de la cercosporiose noire du bananier, a été décrit par différents auteurs. Les études montrent que les conditions climatiques (température et humidité) agissent à toutes les étapes du cycle [\[Jones, 2000\]](#) et que l'intensité d'infection (densité des lésions) et la réceptivité de la plante à la maladie sont des facteurs influents sur le cycle infectieux. Cependant, aucune étude n'a jamais été menée pour expliquer l'effet des facteurs environnementaux sur la dynamique d'invasion de la maladie.

La vitesse d'invasion d'une épidémie dépend des conditions climatiques et de la structure du

paysage [Plantegenest et al., 2007] [Parnell et al., 2010]. En effet, une analyse globale des données historiques (relevées par enquête auprès des ministères de l'agriculture des pays) de l'invasion de la cercosporiose noire dans 13 pays des Caraïbes suggère que la vitesse d'invasion dépend des conditions climatiques, de la structure du paysage (topographie et composition du paysage) et de l'application des méthodes de lutte [Abadie C., données personnelles]. Ainsi, les principaux facteurs environnementaux jouant un rôle dans la propagation spatiale de l'épidémie de la cercosporiose noire du bananier sont d'une part les conditions climatiques et d'autre part la structure du paysage (incluant les méthodes de luttes).

Facteurs climatiques Le vent est nécessaire à la propagation à longue distance des ascospores (spores éjectées des périthèces). En effet, dans chacune des expériences menées pour évaluer la distance de dispersion des ascospores, la direction et la vitesse du vent ont été mesurées et intégrées dans les différents calculs de dispersion. L'humidité ou les précipitations (optimales en conditions climatiques tropicales) sont aussi nécessaires à la germination des spores sur les feuilles de bananier et à la sporulation du champignon [Stover, 1980] [Churchill, 2011] (cf. Chapitre 1).

Structure du paysage Contrairement aux différents cas d'études de la dispersion de *Mycosphaella fijiensis* cités précédemment, à l'échelle d'un territoire la structure du paysage n'est pas homogène. L'invasion du pathogène ne dépend donc pas seulement de la distance entre plantes et des conditions climatiques de la zone d'étude, elle dépend aussi des caractéristiques du paysage par la définition de capacités de réception des cultivars de la zone à envahir [Parnell et al., 2010] [Neri et al., 2011]. Par exemple, un bananier de type Cavendish (banane dessert) est plus sensible à la cercosporiose noire qu'un bananier sauvage ou qu'un bananier plantain [Carlier et al., 2000]. De plus, dans la mesure où le pathogène est spécifique du bananier, on peut émettre l'hypothèse que la probabilité d'invasion est plus faible dans les zones à faible densité de plantation que dans celles à forte densité.

4.1.2 Contexte méthodologique

Une manière commune de modéliser un système épidémique est de considérer que chaque individu progresse suivant différents états épidémiologiques dans le temps et dans l'espace [Gilligan, 2008]. Compte tenu de la nature stochastique de l'évolution d'épidémies de maladies infectieuses, de nombreux modèles stochastiques (les transitions entre différents états sont gouvernés par des probabilités) sont utilisés pour explorer les mécanismes biologiques de propagation d'épidémies dans des espaces homogènes (Fewster [2003], Filipe and Maule [2004], Long et al. [2012], Ross [2011]).

En paysage hétérogène, les modèles stochastiques d'invasions biologiques s'intéressent à comprendre le succès d'une invasion selon différents types d'habitats et expliquer la propagation spa-

tiale et temporelle (Schreiber and Lloyd-Smith [2009], Meentemeyer et al. [2011], Cook et al. [2007]). En effet, d'après Parnell et al. [2010] ou encore Henderson et al. [2006] précisément sur la cercosporiose noire, il semble possible de manipuler l'hétérogénéité du paysage pour aider à l'éradication d'épidémies foliaires.

Lorsque l'on désire étudier les localisations spatiales et temporelles d'occurrences liées à un phénomène, les processus ponctuels sont souvent utilisés car les occurrences sont associés à des mesures stochastiques [Daley and Vere-Jones, 2003]. Il s'agit pour chaque date et chaque localisation spatiale de spécifier le processus conditionnellement à l'histoire passée. Rasmussen et al. [2007] décrivent au moyen d'un processus ponctuel spatio-temporel en temps continu la propagation d'un insecte ravageur sur une grille lattice. Vaillant et al. [2010] utilisent aussi des processus ponctuels pour analyser la propagation du *yellow leaf virus* de la canne à sucre dans un champ et inférer les paramètres du modèle en utilisant des données d'occurrence de la maladie sur toutes les plantes à plusieurs dates.

En effet, l'estimation des paramètres des modèles spatio-temporels est un enjeu crucial pour une meilleure compréhension des phénomènes [Hartig et al., 2011] mais les données disponibles d'incidence de maladie à grande échelle sont rares et très souvent censurées : les temps de colonisation sont inconnus. Dans le cadre de données censurées, de nombreuses méthodes d'inférence bayésiennes ont été développées et plus particulièrement des méthodes permettant de pallier le problème de données non observées (Gibson [1997], O'Neill [2002], Filipe et al. [2004]). De telles censures rendent l'évaluation de la vraisemblance difficile. En introduisant alors des variables latentes pour tenir compte des événements non observés, l'inférence peut être réalisée dans un cadre bayésien en augmentant l'espace des paramètres pour y inclure ces événements non observés (Gibson et al. [2010], Tanner and Wong [2010], Brown et al. [2013]).

4.1.3 Objectifs du chapitre

Dans ce chapitre, nous nous intéressons à l'invasion de la cercosporiose noire à l'échelle de la Martinique. L'objectif est d'évaluer la dispersion à longue distance et de mieux comprendre l'effet du climat et de la couverture du paysage dans la dynamique d'invasion.

Pour cela, nous disposons de données d'incidence (présence-absence) de la maladie référencées, dès le début de l'invasion, dans le temps et dans l'espace, ainsi que de données climatiques correspondant à la période de suivi de la maladie. Afin d'améliorer nos connaissances sur la distance de dispersion de *Mycosphaerella fijiensis* à l'échelle d'un territoire, et afin d'évaluer l'influence des facteurs environnementaux sur l'invasion du pathogène, nous avons développé un modèle stochastique spatio-temporel simulant l'invasion du champignon dans un paysage hétérogène.

Nous présenterons d'abord les données et le modèle spatio-temporel développé, puis nous présenterons l'algorithme utilisé dans un paradigme bayésien pour l'inférence des paramètres du

modèle et sa validation sur un jeu de données simulé. Pour finir, nous présenterons l'évaluation de l'influence des facteurs environnementaux sur la dynamique d'invasion de la cercosporiose noire en utilisant les résultats de deux analyses de sensibilité. L'article en cours de finalisation, inséré en fin de chapitre, présente le modèle ainsi que la méthode d'inférence des paramètres sur les données disponibles censurées.

4.2 Données

4.2.1 Données de présence-absence à l'échelle d'un territoire

La cercosporiose noire a fait l'objet d'une analyse de risque phytosanitaire en 2003, et cette même année *Mycosphaerella fijiensis* a été classé comme organisme nuisible de quarantaine pour les départements d'outre-mer (dont la Martinique). Dans le cadre du projet PANDORE coordonné depuis 2009 par le SPAV (Pôle santé et protection des animaux et des végétaux, du service de l'alimentation de la Direction de l'Agriculture et de la Forêt) et dont l'objectif est une détection précoce et une éradication de maladies nouvelles sur le territoire martiniquais, le parcellaire de bananiers était surveillé de façon régulière vis-à-vis de la cercosporiose noire. Des observations de symptômes sur les bananiers complétées par une collecte d'échantillons foliaires en cas de symptômes douteux ont été réalisées deux fois par mois sur 20 parcelles sentinelles fixes (dans le temps et l'espace) de plantains et 8 parcelles de la variété d'exportation Cavendish, choisies aléatoirement sur le territoire. Au cours de l'année 2010, la surveillance des bananiers dans les jardins privés et sur les bords de route a également été assurée (268 sites essentiellement localisés dans la moitié Sud de la Martinique ont été observés). Le 17/09/2010, la maladie a été détectée officiellement sur une des parcelles sentinelles de plantain du réseau.

Les services de l'état (DAAF) et la FREDON Martinique (Fédération Régionale de Défense contre les Organismes Nuisibles) a réalisé un suivi spatiotemporel de la maladie de septembre 2010 (semaine 39) à mai 2012 (soit une durée totale de 83 semaines). La fin de la surveillance du territoire a été décidée par les services de l'état (DAAF) lorsqu'une détection positive eut été confirmée dans chaque commune administrative.

Pour ce suivi, la Martinique a été divisée en 1236 zones élémentaires de 1 km × 1 km et, dans chacune des zones, les bananiers observés positifs ou négatifs à la cercosporiose noire ont été classés suivant cinq types de système de culture (Figure 4.2).

1. Cavendish (bananes dessert cultivées en monoculture pour l'exportation)
2. plantain (bananes légumes cultivées en monoculture pour le marché local)
3. jardin créole (bananiers cultivés en association avec plusieurs cultures)
4. jardin privé (quelques bananiers chez des particuliers)
5. bananier spontané (bananier isolé le long des routes)

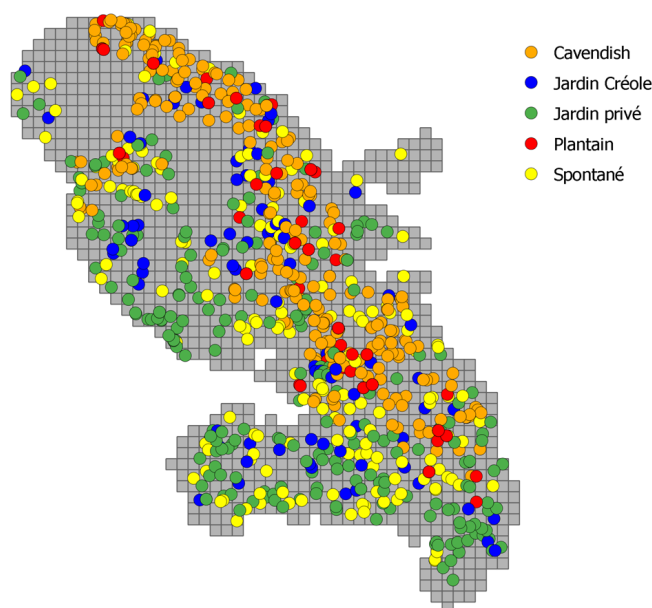


FIGURE 4.2 – Localisation des bananiers observés en 2010 selon leur type de production. La Martinique est discrétisée en 1236 zones.

Tous les bananiers suspectés positifs à la maladie après inspection visuelle ont été échantillonnés et soumis à une analyse en laboratoire (détection moléculaire) permettant de confirmer ou non la présence du pathogène. Sur la période d'invasion, 1635 observations ont été réalisées. A chaque observation sont associés la date (année et semaine), les coordonnées géographiques (longitude et latitude) du bananier, le type de production, et l'incidence de la maladie (présence ou absence).

Les surfaces cultivées et l'importance économique de ces cinq systèmes de culture sont très différentes, par exemple la variété Cavendish pour l'exportation représente la production agricole la plus importante économiquement sur l'île et est cultivée sur 20% de la surface agricole utile. Le suivi hebdomadaire des zones a donc été effectué de manière hétérogène dans le temps et dans l'espace : la majorité des bananiers observés sur les trois années sont de type Cavendish (Figure 4.3), certaines zones n'ont pas été visitées (583 zones soit 47.17 %). Pour les zones visitées au moins deux fois, le temps écoulé entre deux observations est compris entre 1 et 82 semaines avec une moyenne de 21 semaines. Enfin, sur certaines zones, on ne connaît que la semaine de la première détection positive (121 zones soit 9.79 %), ou bien la semaine de la dernière observation négative (364 zones soit 29.45 %). Le passage d'un état négatif à un état positif n'a été observé d'une semaine à la suivante que dans trois zones (0.24% des zones).

La fréquence des observations a rapidement diminué au cours du temps avec la majorité des observations réalisées au cours des trois premiers mois de l'invasion. En effet, 53% du nombre total des observations (865 données) ont été acquises au dernier trimestre 2010 qui correspond à 17% de la durée de l'étude. Ensuite, 40% (654 données) ont été acquises en 2011 et 7% (116 données) en 2012.

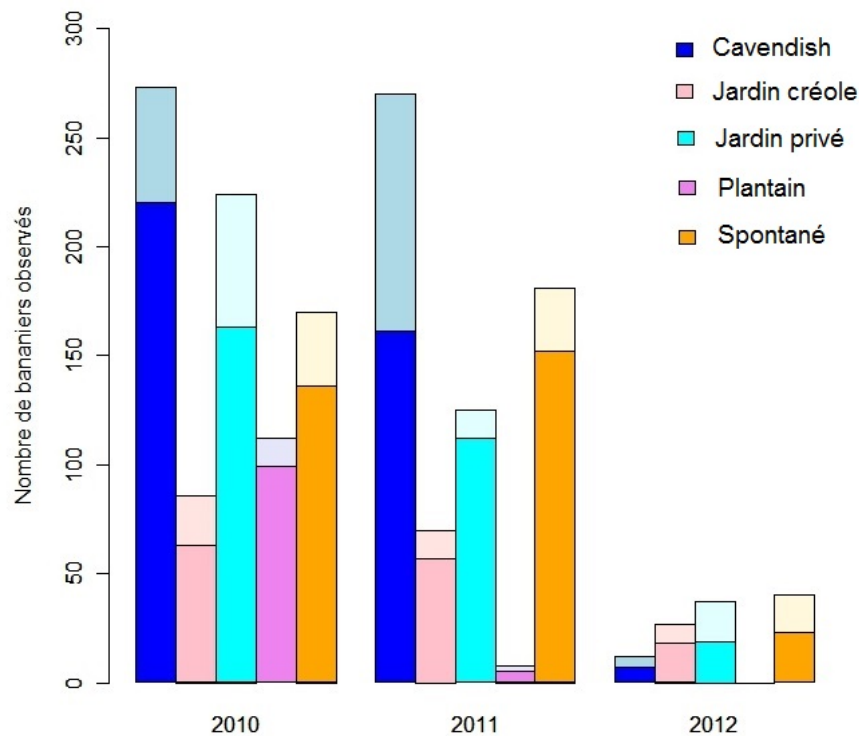


FIGURE 4.3 – Nombre de bananiers observés chaque année de suivi* selon les types de production. La partie claire des barres représente le nombre de bananiers diagnostiqués positifs à la cercosporiose noire, et la partie foncée à ceux diagnostiqués négatifs. *Les années 2010 et 2012 correspondent à une période d'un seul trimestre.

Les données d'incidence disponibles sont donc des données censurées, la colonisation de la zone i (nommée par la suite "patch i ") a été effective dans une fenêtre d'observation $w_i = [\ell_i, v_i]$ où ℓ_i est la semaine de la dernière observation négative et v_i est la semaine de la première observation positive.

4.2.2 Données environnementales

Parcellaire déclaré

Le parcellaire géoréférencé des surfaces déclarées et cultivées en bananiers de type Cavendish et plantain a été fourni par la DAAF pour les années 2010 (Figure 4.4), 2011 et 2012.

La surface cultivée déclarée de plantain est environ vingt fois inférieure à celle de Cavendish. Les surfaces cultivées déclarées sont identiques en 2011 et 2012 pour les deux types de productions (Tableau 4.1).

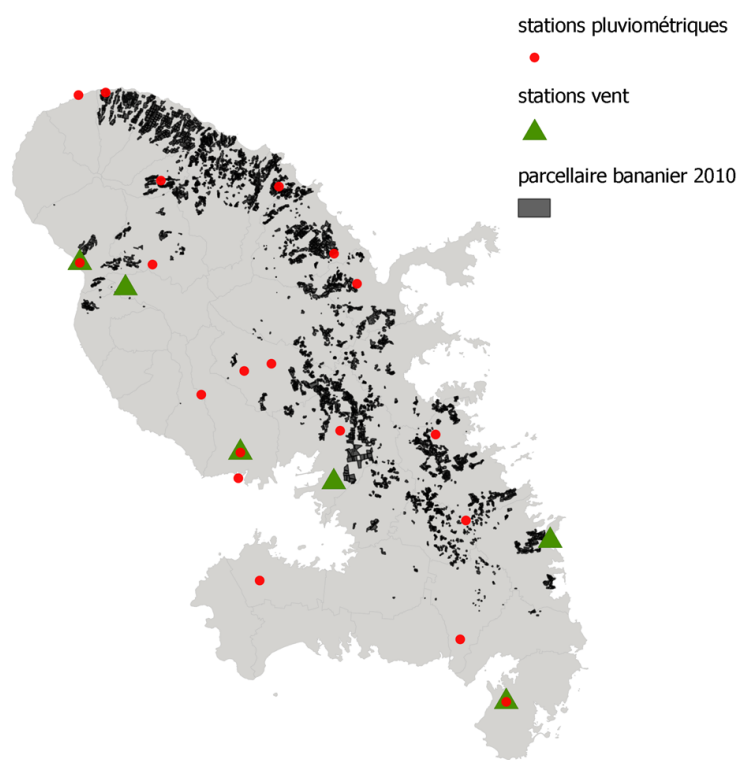


FIGURE 4.4 – Cartographie du parcellaire 2010 des bananes Cavendish et plantain déclaré et des stations météorologiques. Données fournies par la DAAF (parcellaire) et Météo France Martinique (stations météorologiques).

	2010	2011 et 2012
Surface plantain	290.80	302.48
Surface Cavendish	6217.58	6073.37
% de surface occupée	5.51%	5.38 %

TABEAU 4.1 – Surfaces cultivées en bananiers de type Cavendish (dessert pour l'exportation) et plantain (pour le marché locale) déclarées de 2010 à 2012 (hectares).

Pluviométrie et vent

Nous avons acquis de Météo France Martinique les données journalières évaluant la hauteur de pluie (mm), la vitesse (m/s) et la direction (degrés) du vent sur différentes stations météorologiques réparties sur la Martinique (Figure 4.4). Les moyennes hebdomadaires des hauteurs de pluie recueillies sur 19 stations ont été calculées pour l'ensemble de la période de suivi (83 semaines). La variance de ces mesures représente l'effet saisonnier (avec moins de pluies de décembre à avril). Les quantités hebdomadaires moyennes les plus élevées ont été enregistrées dans la moitié Nord de l'île (stations numéro 3, 10 et 11) et les deux quantités les plus faibles ont été enregistrées dans la partie Sud (stations numéro 15 et 19, Figure 4.5). Statistiquement, le test de

Kruskal-Wallis (test non paramétrique permettant de tester si les 19 échantillons proviennent de la même population) nous permet de conclure au seuil de 5% à l'existence d'une variabilité spatiale entre les distributions (Figure 4.5).

Les six stations météorologiques permettant d'obtenir des informations sur la vitesse et direction du vent n'étant pas réparties de façon homogène sur le territoire et sachant un réel gradient entre la côte Est dite "côte au vent" et la côte Ouest dite "côte sous le vent", nous ne considérerons pas par la suite ces données.

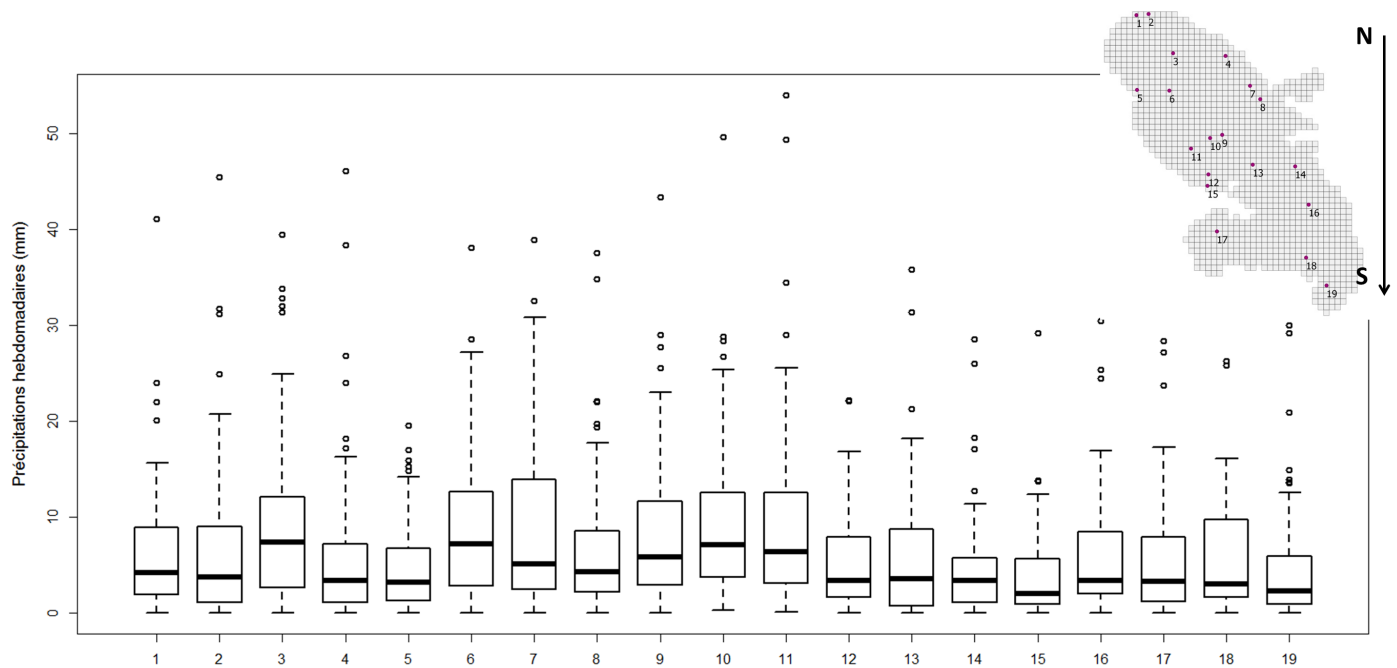


FIGURE 4.5 – Évaluation des quantités hebdomadaires de pluies mesurées sur les 85 semaines de l'étude, selon un gradient spatial Nord/Sud sur les 19 stations météorologiques disponibles.

4.2.3 Reconstitution du paysage

Afin d'intégrer toutes les informations spatio-temporelles disponibles, nous avons renseigné chacun des 1236 patchs en utilisant les données environnementales et en se basant sur la description des cinq types de production déjà disponibles.

Tout d'abord, nous avons attribué à chaque patch la proportion de couverture du paysage de chacun des cinq types de production (Figure 4.6).

Les proportions des types de production Cavendish ou plantain ont été calculées grâce au parcellaire déclaré disponible (Figure 4.4) et au système d'information géographique QGIS (version 2.8). Nous remarquons que la répartition spatiale de ces deux types de production n'est pas homogène sur le territoire, ces deux productions étant essentiellement concentrée sur la moitié Nord de l'île (la plus pluvieuse) et pour la majorité sur la côte Est (Figures 4.4 et 4.6). Pour les trois autres types de production, nous nous sommes appuyés sur des opinions d'experts présents en Marti-

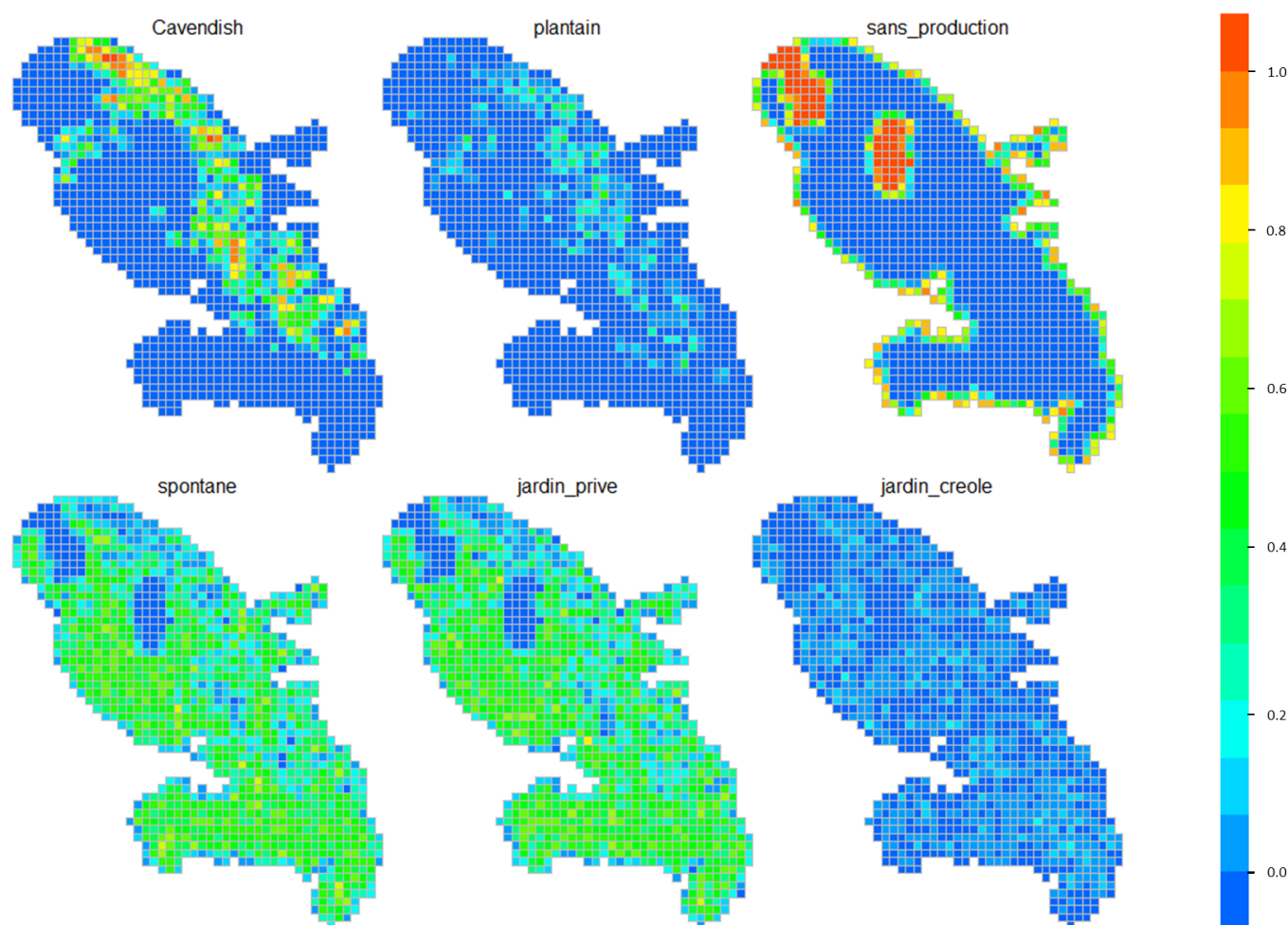


FIGURE 4.6 – Proportions moyennes de couverture du paysage par les six types de production dans chacun des 1236 patchs. Les proportions ont été moyennées sur les 83 semaines d'étude.

nique selon lesquels le type de production jardin créole est rare sur le territoire tout comme le type plantain. Les jardins privés et bananiers spontanés sont quant à eux majoritaires et présents sur tout le territoire. Enfin, nous avons attribué le type "sans production" aux zones paysagères où aucun bananier n'est présent (volcan/mer).

Ensuite, pour estimer la pluviométrie des 1236 patchs, nous avons attribué à chacun d'eux la pluviométrie de la station météorologique la plus proche, ce qui a permis de diviser la Martinique en 19 zones pluviométriques (Figure 4.7).

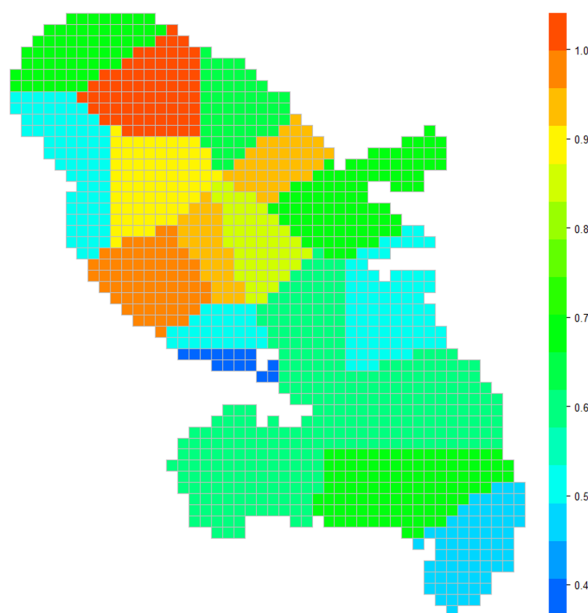


FIGURE 4.7 – Quantités moyennes de pluies dans les 19 zones correspondantes aux 19 stations météorologiques. Les quantités pluviométriques ont été moyennées sur les 83 semaines d'étude et transformées de telle sorte qu'elles soient comprises entre 0 et 1.

4.3 Modèle spatio-temporel

Nous avons développé un modèle spatio-temporel stochastique basé sur des processus ponctuels en temps continu et espaces discrétisés intégrant la structure du paysage. Cette dernière est prise en compte en considérant des types de cultures et des hauteurs pluviométriques associées à chaque patch traduisant la capacité du patch à propager l'épidémie. Les notions de processus ponctuel ont été rappelées dans le Chapitre 1.

Notre approche est similaire à celle développée par [Cook et al. \[2007\]](#) qui modélisent à l'échelle d'un pays la distribution spatiale de l'extension territoriale d'une plante invasive *Heracleum mantegazzianum* en utilisant des covariables qui prennent en compte la structure du paysage.

4.3.1 Description du modèle développé

Nous avons développé un modèle qui tient compte de la capacité de réception des spores d'un type d'habitat. Les notations utilisées figurent dans le tableau 4.2.

La dispersion du pathogène est modélisée par une fonction de dispersion $f_{\beta}(d_{i,j})$ qui dépend de la distance $d_{i,j}$. Cette fonction peut prendre plusieurs formes [[Austerlitz et al., 2004](#)]. Dans notre cas d'étude, il importe cependant de choisir une fonction de dispersion qui permette de la dispersion à longue distance, souvent caractérisée par des fonctions dites "à queue lourde". La fonction choisie devra être aussi isotrope englobant alors l'hypothèse d'une propagation de la maladie par des modes de dispersion de type anthropique (transport par l'homme de matériel végétal infecté) ou de type naturel (transport de spores par le vent).

Soit $\lambda(x, t)$ pour $x \in X = \{1, \dots, N\}$ l'intensité du processus ponctuel spatio-temporel à l'instant t . Soit $I_{t^-} = \{x \in S | \tau_x < t\}$ l'ensemble des patches x infectés avant la date t (non incluse), on définit l'intensité conditionnelle du processus :

$$\lambda(j, t | \mathcal{F}_{t^-}) = \left[R_j \sum_{i \in I_{t^-}} f_{\beta}(d_{i,j}) \right] \mathbf{1}_{\{j \in C^* \setminus I_{t^-}\}} \quad (4.1)$$

où \mathcal{F}_{t^-} est l'histoire du processus jusqu'à la date t exclue et $f_{\beta}(d_{i,j})$ la fonction de dispersion dépendant de la distance euclidienne $d_{i,j}$ entre le patch i et le patch j et du paramètre β .

La capacité R_j du patch j à recevoir des spores est décrite par des covariables que l'on suppose influentes : la pluviométrie et les types de production qui composent le patch.

$$R_j = \exp(\gamma \mathbf{P}_j) \sum_{k=0}^5 (r_k \mathbf{H}_{j,k}) \quad (4.2)$$

où \mathbf{P}_j est la hauteur de pluie moyenne dans le patch j et $\gamma \in (-\infty, \infty)$ est un paramètre représentant l'effet de la pluie dans la capacité de réception. On suppose ici que la pluviométrie intervient de manière exponentielle dans la réceptivité d'un patch à la maladie.

$\mathbf{H}_{j,k}$ est la proportion de couverture du patch j par le type de production k , $k \in \{0, \dots, 5\}$ et $r_k \in [0, \infty)$ sont des paramètres représentant l'aptitude du type de production k à être colonisé (les types de production sont respectivement,

- 0 : sans production,
- 1 : bananier spontané
- 2 : jardin privé,
- 3 : jardin créole,
- 4 : Cavendish,
- 5 : plantain.

On pose $r_0 = 0$ pour tenir compte du fait que le type sans production est non colonisable (volcan/mer).

Le vecteur des paramètres de ce modèle est $\Theta = (\beta, \gamma, r_0, \dots, r_5)$.

Nous utiliserons une fonction de dispersion isotrope de type puissance, de la forme $f_{\beta}(d_{i,j}) = \beta^2 d^{-2\beta}$. Cette fonction a queue lourde est caractérisée par des événements de colonisation à longue distance plus nombreux que ceux obtenus avec une fonction de dispersion exponentielle (Mundt and Sackett [2012], Hastings et al. [2005]) ce qui semble être bien adapté à notre cas d'étude.

Notation	Définition
X	$= \{1, \dots, N\}$ ensemble des patches avec N le nombre de patches total en Martinique
D	$= \{1, \dots, T\}$ durée de l'étude avec T dernière semaine de l'étude
\mathcal{F}_t^-	histoire du processus jusqu'à la date t exclue
τ_i	temps de colonisation du $i^{\text{ème}}$ patch (en semaine suivant le début de l'étude)
τ	$= \{\tau_1, \dots, \tau_N\}$ ensemble des temps de colonisation des patches
τ_{end}	temps de la dernière observation
C	ensemble des patches colonisés avant τ_{end}
C^*	ensemble des patches non-colonisés avant τ_{end}
I_{t^-}	l'ensemble des patches x infectés avant la date t .
$d_{i,j}$	distance entre deux patches i et j (km)
\mathcal{N}_i	$= \{j : d_{i,j} \leq d_{max}\}$, le voisinage de i à une distance d_{max}
Θ	$= (\beta, \gamma, r_1, r_2, r_3, r_4, r_5)$, vecteur des paramètres du modèle
ℓ_i	temps de dernière observation négative dans le patch i (semaine)
v_i	temps de première observation de la maladie dans le patch i (semaine)
w_i	$= (\ell_i, v_i)$ fenêtre d'observation du patch i
W	ensemble des observations censurées i.e. $W = \{w_i : i = 1, \dots, N\}$
\mathbb{X}_I	ensemble de patches responsables de l'inoculum primaire (patches infectés au début de l'étude)
\mathbb{X}_M	ensemble de patches non colonisables
$\lambda(i, t \mathcal{F}_{t^-})$	l'intensité du processus ponctuel d'un patch i au temps t conditionnellement à l'histoire du processus avant t \mathcal{F}_{t^-} (notation : $\lambda(i, t)$ pour plus de brièveté)
R_j	réceptivité du patch j
P_j	hauteur de pluie moyenne dans le patch j
$H_{j,k}$	proportion de couverture du patch j par le type de production k , $k \in \{0, \dots, 5\}$.
$f_\beta(d_{i,j})$	fonction de dispersion, dépendante de la distance $d_{i,j}$ entre les patches i et j .
$L(\Theta \tau)$	vraisemblance, conditionnelle à \mathbb{X}_I et \mathbb{X}_M (omis dans l'écriture pour plus de brièveté)
$\pi(\Theta \tau)$	loi jointe a posteriori des paramètres
$\pi(\Theta)$	loi jointe a priori des paramètres
$\pi(\beta)$	loi a priori marginale du paramètre β (idem pour les autres paramètres)

TABLEAU 4.2 – Notations utilisées dans le chapitre

4.3.2 Simulations du modèle

Le processus considéré étant un processus ponctuel spatio-temporel d'intensité $\lambda(x, t)$ avec $x \in X$, $t \in D$, conditionnellement à l'histoire \mathcal{F}_{t^-} , le temps d'attente est distribué selon une loi exponentielle de paramètre $\lambda(x, t)$ et le temps d'attente de la prochaine infection suit une loi expo-

nentielle de paramètre $\sum_{x=1}^N \lambda(x, t)$. Il est utile de préciser que $\lambda(x, t) = 0$ si x est déjà infecté à la date t .

La simulation se décompose en deux étapes : d'abord générer la date de la future infection puis établir la position du patch infecté à cette date (Algorithme 3).

Étape 1 : calculer le temps d'attente de la prochaine colonisation.

Le temps d'attente du prochain évènement est

$$\epsilon = \inf(\tau_x) \sim \exp\left(\sum_{x=1}^N \lambda(x, t)\right) \quad (4.3)$$

Étape 2 : établir la position du patch concerné par le temps de future colonisation ($t + \epsilon$).

Le patch j est retenu avec probabilité

$$p_j = \frac{\lambda(j, t)}{\sum_{j=1}^N \lambda(j, t)} \quad (4.4)$$

Algorithme 3 Processus ponctuel spatio-temporel sur $[1, N] \times [1, T]$

1. Initialisation

$t = 0$

2. Étapes 1 & 2

tant que $t < T$ **faire**

calculer $\sum_{x=1}^N \lambda(x, t)$

tirer $\epsilon \sim \exp(\sum_{x=1}^N \lambda(x, t))$

$t \leftarrow t + \epsilon$

tant que $t < T$ **faire**

tirer $u \sim \mathcal{U}[0, 1]$

pour $j = 1$ **faire**

calculer $\sum_{x=1}^j \left[\frac{\lambda(x, t)}{\sum_{x=1}^N \lambda(x, t)} \right]$

si $u < \sum_{x=1}^j \left[\frac{\lambda(x, t)}{\sum_{x=1}^N \lambda(x, t)} \right]$ **alors**

$\tau_j = t$

fin du pour

sinon

$j = j + 1$

fin du si

fin du pour

fin du tant que

fin du tant que

Paramétrage du modèle. Pour réaliser des simulations réalistes de l'invasion de la cercosporiose noire en Martinique (Figures 4.8 et 4.9) nous avons choisi les paramètres décrivant l'influence des types de production en utilisant des informations d'expertises (de la DAAF et/ou du CIRAD) pour

chaque type de production : les densités de peuplement, les variétés cultivées (qui présentent des niveaux de sensibilité à la maladie différents) mais aussi l'utilisation de fongicide pour lutter contre le pathogène (Tableau 4.3).

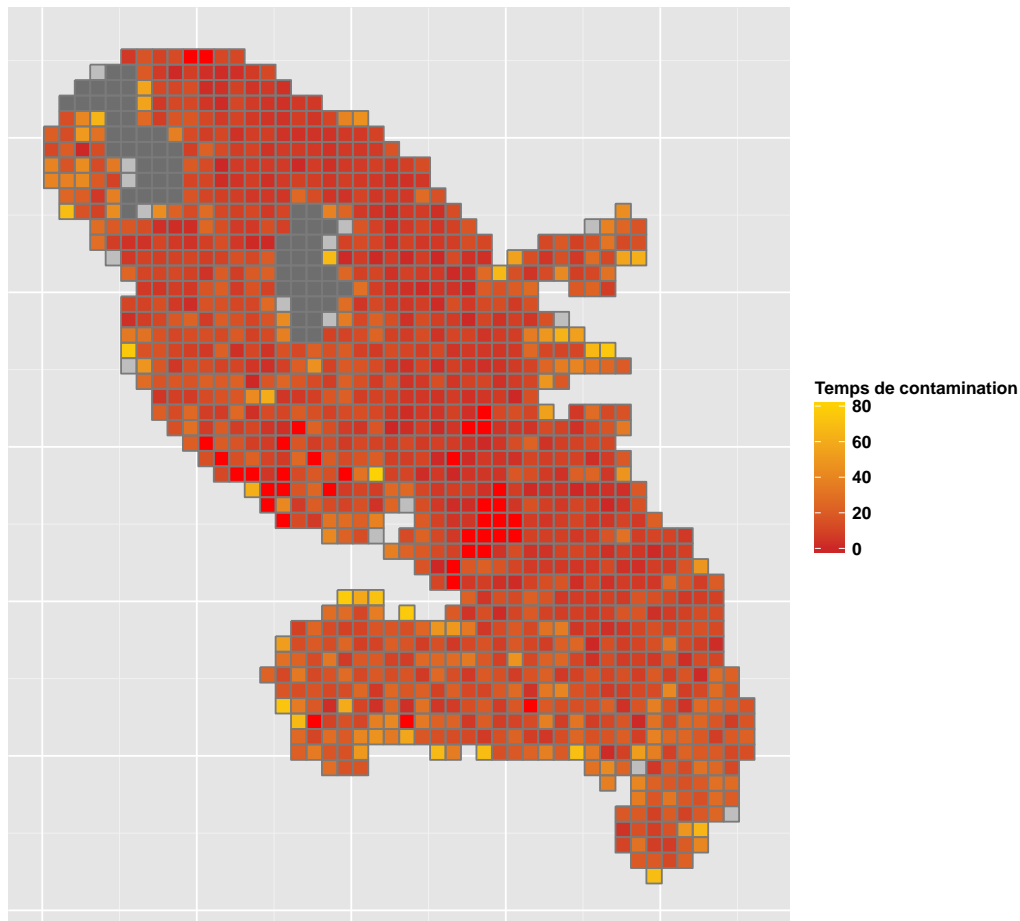


FIGURE 4.8 – Simulation d'une épidémie de cercosporiose noire à l'échelle de la Martinique avec zones de premières infections réelles (zones rouges) sur 1236 patchs avec les paramètres $\beta = 0.4$, $\gamma = 0.3$, $r_1 = 0.002$, $r_2 = 0.005$, $r_3 = 0.008$, $r_4 = 0.04$, $r_5 = 0.03$. Les patchs gris foncé sont les patchs non colonisables, les patchs gris clair sont les patchs non colonisés.

Les bananiers de type Cavendish sont les plus sensibles à la maladie et sont ceux cultivés principalement pour l'exportation. Ces cultures sont donc traitées par fongicide, ce qui diminue fortement leur capacité de production de spores, mais leur forte densité de plantation (1800 plants/hectares) les rendent très réceptives à l'invasion.

Le type de production plantain est également sensible à la maladie (cf. chapitre 1) et a la même densité de plantation que le type Cavendish. Le type plantain étant destiné au commerce local, l'application de fongicide n'est pas réalisée.

Les jardins créoles et privés sont composés de plusieurs types de cultures (bananières, maraîchères, autres fruits) et ont une densité de plantation beaucoup moins élevée que celle du type plantain ou Cavendish : les jardins créoles sont de petites exploitations de quelques hectares (10

à 30 plants) contrairement aux jardins privés qui ne représentent que quelques plants. Les jardins créoles ont donc une réceptivité à la maladie plus élevées que les jardins privés.

Les bananiers de type spontané sont des bananiers de bord de route et appartiennent souvent à des variétés partiellement résistantes : ils ont donc une faible réceptivité à la maladie.

Réceptivité à la maladie	
très élevée	Cavendish traitée plantain jardin créole jardin privé
très faible	Spontané

TABLEAU 4.3 – Classement des capacités de réceptivité à la maladie des différents types de productions.

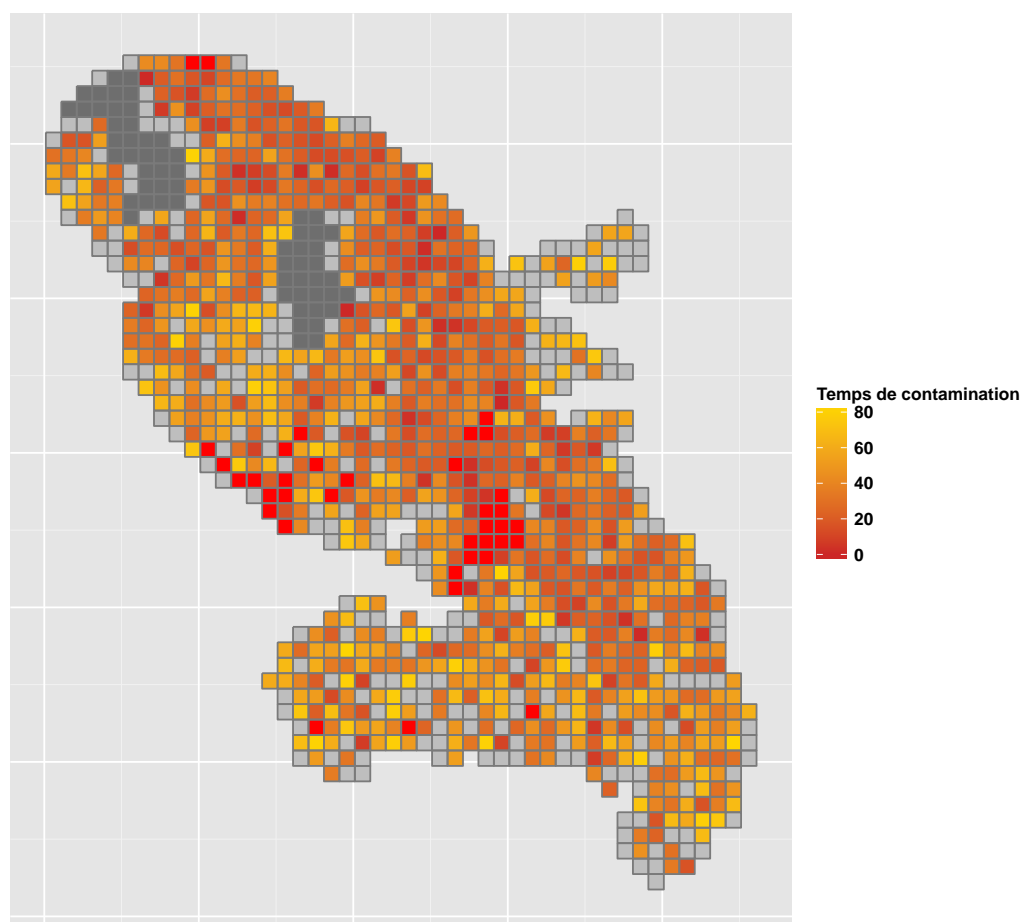


FIGURE 4.9 – Simulation d'une épidémie de cercosporiose noire à l'échelle de la Martinique avec zones de premières infections réelles (zones rouges) sur 1236 patches avec les paramètres $\beta = 0.1$, $\gamma = 0.1$, $r_1 = 0.002$, $r_2 = 0.005$, $r_3 = 0.008$, $r_4 = 0.04$, $r_5 = 0.03$. Les patches gris foncé sont les patches non colonisables, les patches gris clair sont les patches non colonisés.

4.4 Inférence des paramètres du modèle

Nous utilisons une approche Bayésienne pour inférer les paramètres du modèle, et trouver les distributions a posteriori des paramètres sachant les données observées (cf. [Chapitre 1](#)).

Les données utilisées pour l'inférence sont d'une part la matrice $W = \{w_i\}$ où $w_i = [\ell_i, v_i]$ est la fenêtre d'observation des symptômes dans le patch i , mais aussi l'ensemble des patches infectés initialement \mathbb{X}_I et l'ensemble des patches non colonisables \mathbb{X}_M .

4.4.1 Vraisemblance, loi a priori et loi a posteriori

La vraisemblance s'écrit en considérant le processus de colonisation comme un processus ponctuel spatio-temporel en temps continu markovien.

Si le vecteur τ des temps de première colonisation de chaque patch est observé, la fonction de vraisemblance est alors fournie par la théorie des processus ponctuels (voir par exemple [Cressie and Wikle \[2011\]](#)) et s'écrit :

$$L(\Theta|\tau) = \prod_{i \in C} \left[\lambda(i, \tau_i) \times \exp \left\{ - \int_0^{\tau_i} \lambda(i, t) dt \right\} \right] \quad (4.5)$$

$$\times \prod_{i \in C^*} \exp \left\{ - \int_0^T \lambda(i, t) dt \right\}$$

$$L(\Theta|\tau) = \prod_{i \in C} \left[\lambda(i, \tau_i) \times \exp \left\{ - \sum_{j: \tau_j < \tau_i} \lambda(i, t)(\tau_j - \tau_i) \right\} \right] \quad (4.6)$$

$$\times \prod_{i \in C^*} \exp \left\{ - \sum_{j: \tau_j < \tau_{end}} \lambda(i, t)(\tau_{end} - \tau_j) \right\}$$

En l'absence d'information sur les paramètres, les lois marginales a priori des paramètres utilisées sont des loi uniformes largement bornées : $\pi(\beta) \sim U[0, 5]$, $\pi(\gamma) \sim [-10, 10]$, $\pi(r_1) \sim [0, 10]$, $\pi(r_2) \sim [0, 10]$, $\pi(r_3) \sim [0, 10]$, $\pi(r_4) \sim [0, 10]$, $\pi(r_5) \sim [0, 10]$.

Sachant que la loi a posteriori est proportionnelle au produit de la vraisemblance et des lois a priori

$$\pi(\Theta|\tau) \propto L(\Theta|\tau) \pi(\Theta) \quad (4.7)$$

et en considérant des priors indépendants pour chaque paramètre, la loi a posteriori s'écrit :

$$\pi(\Theta|\tau) \propto L(\Theta|\tau) \pi(\beta) \pi(\gamma) \pi(r_1) \dots \pi(r_5) \quad (4.8)$$

4.4.2 Algorithmes et validation

Dans une démarche "pas-à-pas" nous avons dans un premier temps utilisé et validé un algorithme MCMC d'Hastings-Métropolis en utilisant des données non censurées, i.e. des données où les temps de colonisation sont connus [Cook et al., 2007]. Dans un second temps nous avons utilisé et validé un algorithme MCMC d'Hastings-Metropolis avec augmentation de données qui permet d'inférer les paramètres du modèle dans le cas de données censurées, i.e. des données où les temps de colonisation exacts des patches sont inconnus mais où on dispose d'intervalles de temps pendant lesquels les colonisations ont eu lieu (Catterall et al. [2012] Brown et al. [2013]). Pour des raisons de temps de calcul, la validation des algorithmes a été effectuée sur 500 patches (zone Nord de la Martinique).

Les deux algorithmes implémentés en C++ sont présentés et validés sur des jeux de données simulés. La validation des algorithmes s'est faite sur trois jeux de données avec paramétrages différents et, le modèle étant stochastique, sur trois simulations du processus. Dans chacun des cas, la convergence a été établie d'une part visuellement par le tracé des différentes chaînes de Markov des paramètres et au moyen du calcul du critère de Gelman-Rubin \hat{R} [Gelman and Rubin, 1992] sur trois chaînes pour chaque paramètre. Ce critère permet de s'assurer que la phase ergodique des différentes chaînes de Markov est atteinte en calculant un rapport de variance inter et intra-chaîne [Robert and Casella, 2010].

Algorithmes

Des algorithmes bayésiens MCMC on déjà été développés pour inférer des paramètres des processus markovien de colonisation-dispersion en temps continu et espace discrétisé avec des données d'incidence.

L'algorithme 4 est celui proposé par Cook et al. [2007] dans le cas de données non censurées. Les lois de propositions utilisées sont des loi normales centrées sur les valeurs des paramètres des itérations précédentes.

L'algorithme 5 utilise une technique d'augmentation de données utilisée dans Catterall et al. [2012] et Brown et al. [2013]. Dans cet algorithme, les données utilisées sont censurées et seul des intervalles de temps de colonisation sont disponibles. Les temps de colonisation $\tau_i \in [\ell_i, \nu_i]$ sont considérés comme des paramètres supplémentaires et sont inférés de la même manière que les paramètres Θ . Le rapport ρ d'Hastings-Metropolis est calculé en utilisant seulement la vraisemblance et la loi de proposition. Les lois de propositions utilisées pour ces paramètres supplémentaires sont des loi normales tronquées sur l'intervalle w_i de chaque patch.

Algorithme 4 Algorithme MCMC pour données non censurées .

Entrée(s) Lire les données i.e. les temps de colonisation de chacun des patches.

1. Initialisation

Choisir le nombre d'itérations M.

Initialisation du vecteur des paramètres : $\Theta^{[0]} = \{\theta_0^{[0]}, \dots, \theta_k^{[0]}\} \sim \pi(\Theta)$

Calcul de $L(\Theta^{[0]}|\tau)$.

2. Itérations

$m = 1$

tant que $m \leq M$ **faire**

pour $k \in \{1, \dots, 7\}$ **faire**

$\theta^* \sim q(\cdot | \theta_k^{[m-1]})$

$$\rho = \frac{L(\Theta^*) \pi(\theta^*) q(\theta^{(m)}, \theta^*)}{L(\Theta) \pi(\theta) q(\theta^*, \theta^{(m)})}$$

$u \sim \mathcal{U}([0, 1])$

si $u \leq \rho$ **alors**

$\theta_k^{[m]} = \theta_k^*$

sinon

$\theta_k^{[m]} = \theta_k^{[m-1]}$

fin du si

fin du pour

$m = m + 1$

fin du tant que

Sortie(s) $\Theta = (\theta_{m,k})_{\substack{1 \leq m \leq M \\ 1 \leq k \leq 7}}$

Algorithme 5 Algorithme MCMC pour données censurées.

Entrée(s) Lire les données i.e. les intervalles de temps de colonisation de chacun des patches.

1. Initialisation

Choisir le nombre d'itérations M.

Initialisation du vecteur des paramètres : $\Theta^{[0]} = \{\theta_0^{[0]}, \dots, \theta_k^{[0]}\} \sim \pi(\Theta)$ Initialisation des temps de colonisation des patches i : $\tau_i^{[0]} \sim U[\ell_i, \nu_i]$ Calcul de $L(\Theta^{[0]}|W)$.

2. Itérations

 $m = 1$ **tant que** $m \leq M$ **faire** **pour** $k \in \{1, \dots, 7\}$ **faire** $\theta_k^* \sim q(\cdot | \theta_k^{[m-1]})$

$$\rho = \frac{L(\Theta^*) \pi(\theta^*) q(\theta^{(m)}, \theta^*)}{L(\Theta) \pi(\theta) q(\theta^*, \theta^{(m)})}$$

 $u \sim \mathcal{U}([0, 1])$ **si** $u \leq \rho$ **alors** $\theta_k^{[m]} = \theta_k^*$ **sinon** $\theta_k^{[m]} = \theta_k^{[m-1]}$ **fin du si** **fin du pour** **pour** $p \in W$ **faire** $\tau_p^* \sim q(\cdot | \tau_k^{[m-1]})$

$$\rho = \frac{L(\Theta^*) q(\tau^{(m)}, \tau^*)}{L(\Theta) q(\tau^*, \tau^{(m)})}$$

 $u \sim \mathcal{U}([0, 1])$ **si** $u \leq \rho$ **alors** $\tau_k^{[m]} = \tau_k^*$ **sinon** $\tau_k^{[m]} = \tau_k^{[m-1]}$ **fin du si** **fin du pour** $m = m + 1$ **fin du tant que****Sortie(s)** $\Theta = (\theta_{m,k})_{\substack{1 \leq m \leq M \\ 1 \leq k \leq 7}}$

4.4.3 Validation

Les algorithmes 4 et 5 ont été validés sur trois jeux de données simulés (trois fois) avec des valeurs de paramètres différentes.

Considérons ici un jeu de données d'invasion de la cercosporiose noire simulé suivant l'algorithme 3 sur une grille de 500 patches (Figure 4.10). Les paramètres utilisés pour simuler le processus sont les suivants : $\beta = 1$, $\gamma = 0$, $r_1 = 0.002$, $r_2 = 0.008$, $r_3 = 0.01$, $r_4 = 0.05$, $r_5 = 0.03$.

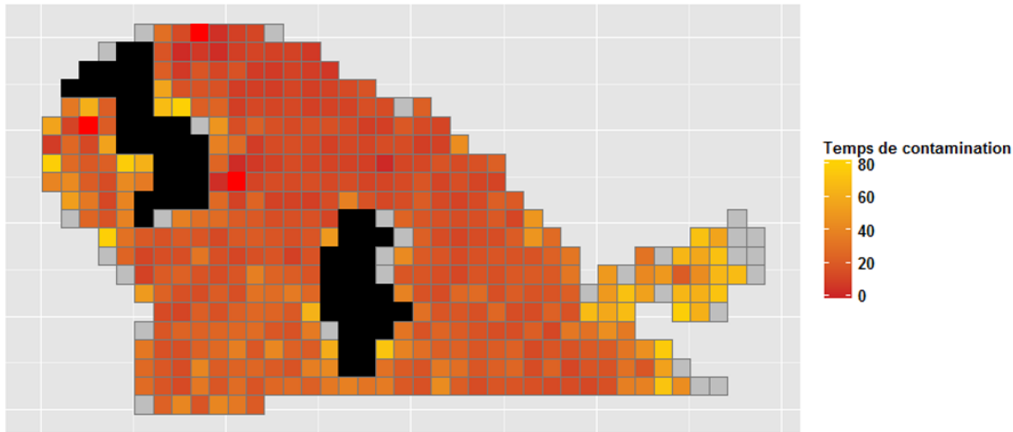


FIGURE 4.10 – Simulation d'une épidémie de cercosporiose noire avec 3 patches infectés en début d'épidémie (en rouge) sur 500 patches (Nord de la Martinique) avec les paramètres $\beta = 1$, $\gamma = 0$, $r_1 = 0.002$, $r_2 = 0.008$, $r_3 = 0.01$, $r_4 = 0.05$, $r_5 = 0.03$. Les patches noirs sont les patches non colonisables, les patches gris clair sont les patches non colonisés.

Théoriquement, une chaîne convergera vers la loi a posteriori. Pour une approximation acceptable de la loi a posteriori, la longueur de la chaîne doit être suffisante pour permettre l'exploration de la densité a posteriori. Cette longueur peut être réduite en faisant en sorte que la chaîne mixe rapidement (en jouant sur les variance des lois de proposition par exemple pour obtenir un taux d'acceptation recommandé [Gelman et al., 2004]). La chaîne ne doit pas être influencée par le choix de la valeur de départ et pour cela, une période dite de "burn-in" doit être définie (elle peut être de longueur variable) et ôtée de l'estimation de la loi a posteriori. Afin de réduire l'autocorrélation des chaînes, il est aussi recommandé de réduire leurs longueurs en effectuant un "thinning" (sélectionner une valeur de la chaîne toutes les x itérations).

La Figure 4.11 propose un diagnostic visuel de convergence des chaînes des 7 paramètres obtenues avec l'algorithme 4 en utilisant les données avec temps de colonisation exacts. Ces traces suggèrent la convergence vers la distribution stationnaire de chacun des paramètres. Le critère de Gelmann et Rubin à aussi été utilisé pour confirmer la convergence : pour toutes les chaînes $\hat{R} = 1$. Les valeurs des paramètres utilisés dans les simulations sont situées au voisinage du mode des densités a posteriori (Figure 4.12). Afin de garantir un taux d'acceptation proche de 25%, les variances des lois de propositions des paramètres de Θ ont été choisies respectivement égales à 0.08, 0.25, 0.009, 0.009, 0.03, 0.03, 0.03.

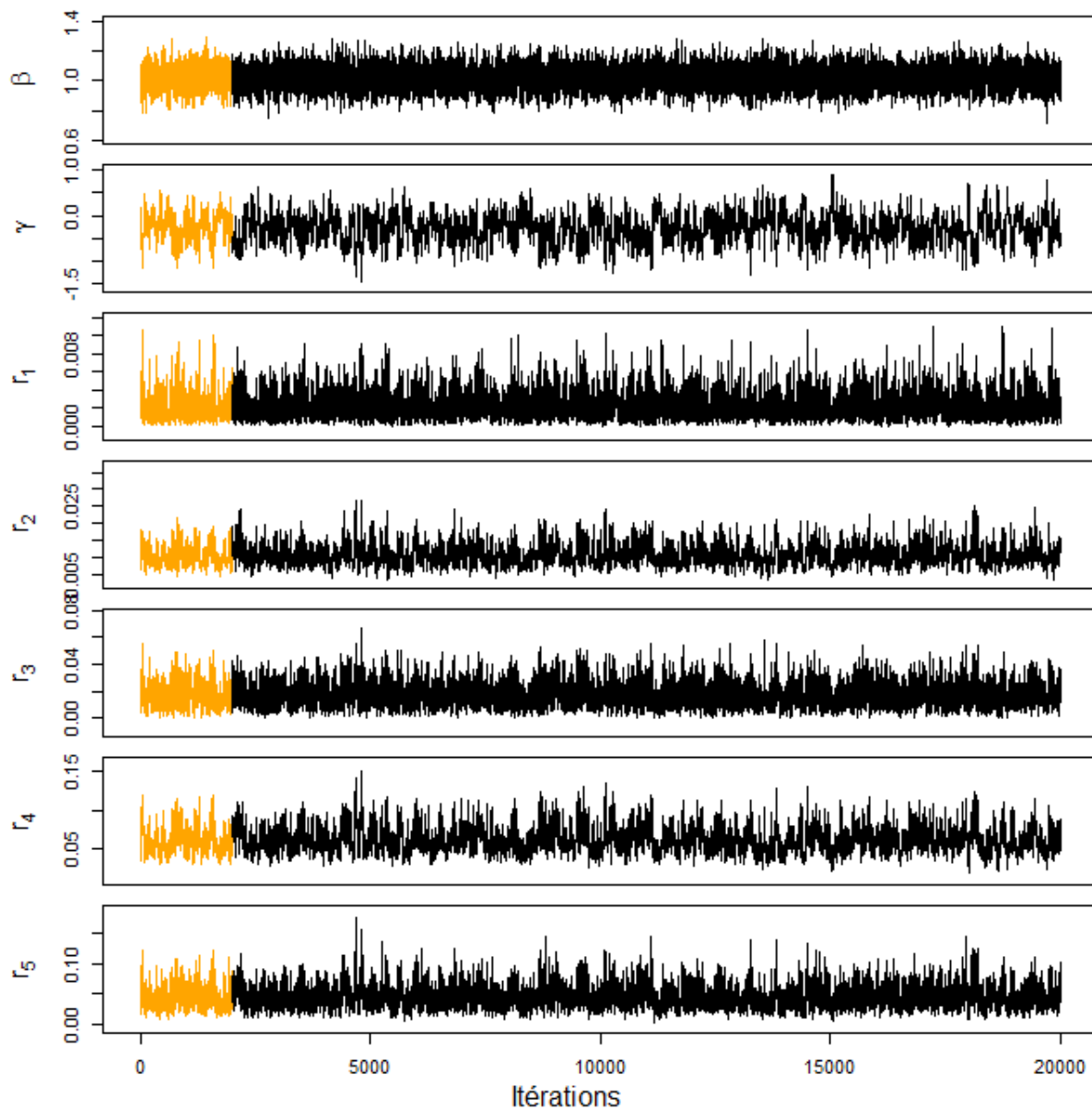


FIGURE 4.11 – Chaînes de Markov des 7 paramètres obtenues avec l'algorithme 4 pour données simulées non censurées. Pour réduire l'autocorrélation des chaînes, nous avons utilisé un "thinning" de 5. La partie orange des chaînes correspond à la période de "burn-in" (2000 itérations), elle ne sera pas utilisée dans les calculs.

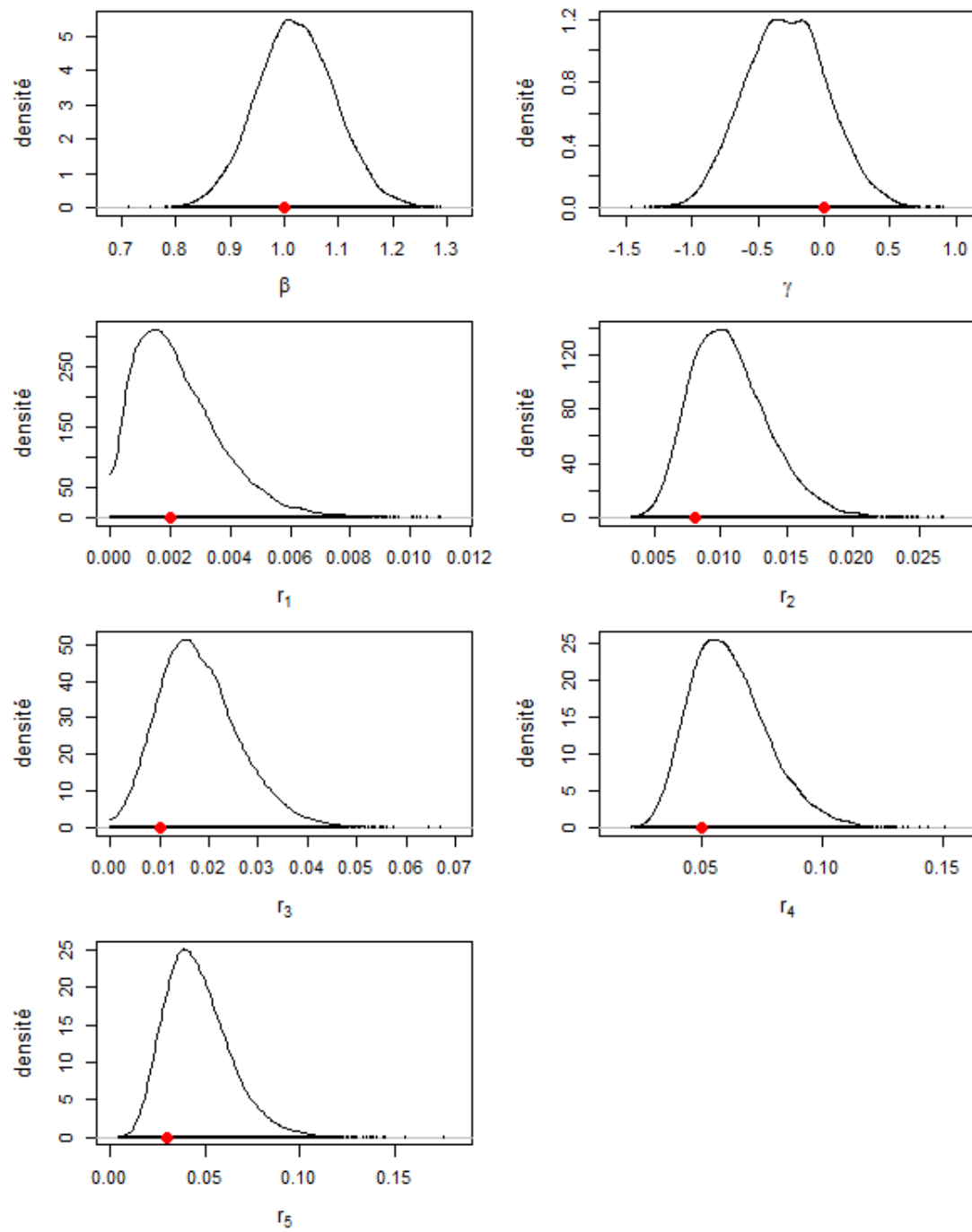


FIGURE 4.12 – Densités marginales a posteriori des 7 paramètres sachant les données simulées non censurées. Les vraies valeurs utilisées pour générer la simulation sont indiquées par des cercles rouges.

La Figure 4.13 propose un diagnostic visuel de convergence des chaînes obtenues avec l'algorithme 5 en utilisant les données censurées en intervalles et la figure 4.14 propose les densités a posteriori des paramètres. Afin de créer un jeu de données censurées, nous avons aléatoirement créé des intervalles autour de chaque temps de colonisation connu en reproduisant au mieux le jeu de données réel i.e. comportant très peu de patchs avec des observations d'une semaine à la suivante, des patchs ayant seulement la dernière observation négative, des patchs ayant seulement une détection positive et des patchs non renseignés.

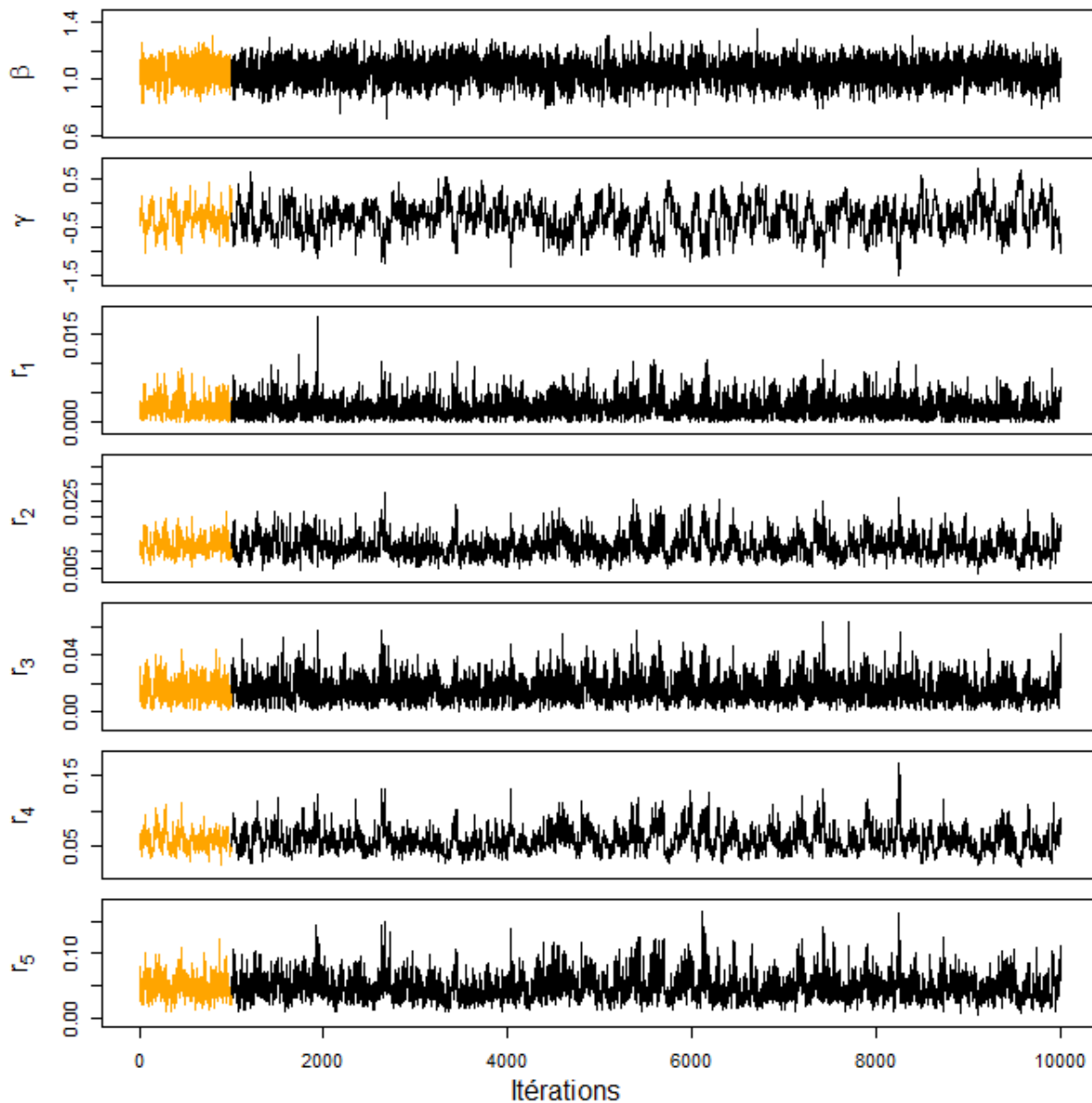


FIGURE 4.13 – Chaînes de Markov des 7 paramètres obtenues avec l'algorithme 5 pour données simulées censurées. Pour réduire l'autocorrélation des chaînes, nous avons utilisé un "thinning" de 5. La partie orange des chaînes correspond à la période de "burn-in" (1000 itérations), elle ne sera pas utilisée dans les calculs.

Ces traces suggèrent la convergence vers la distribution stationnaire de chacun des paramètres.

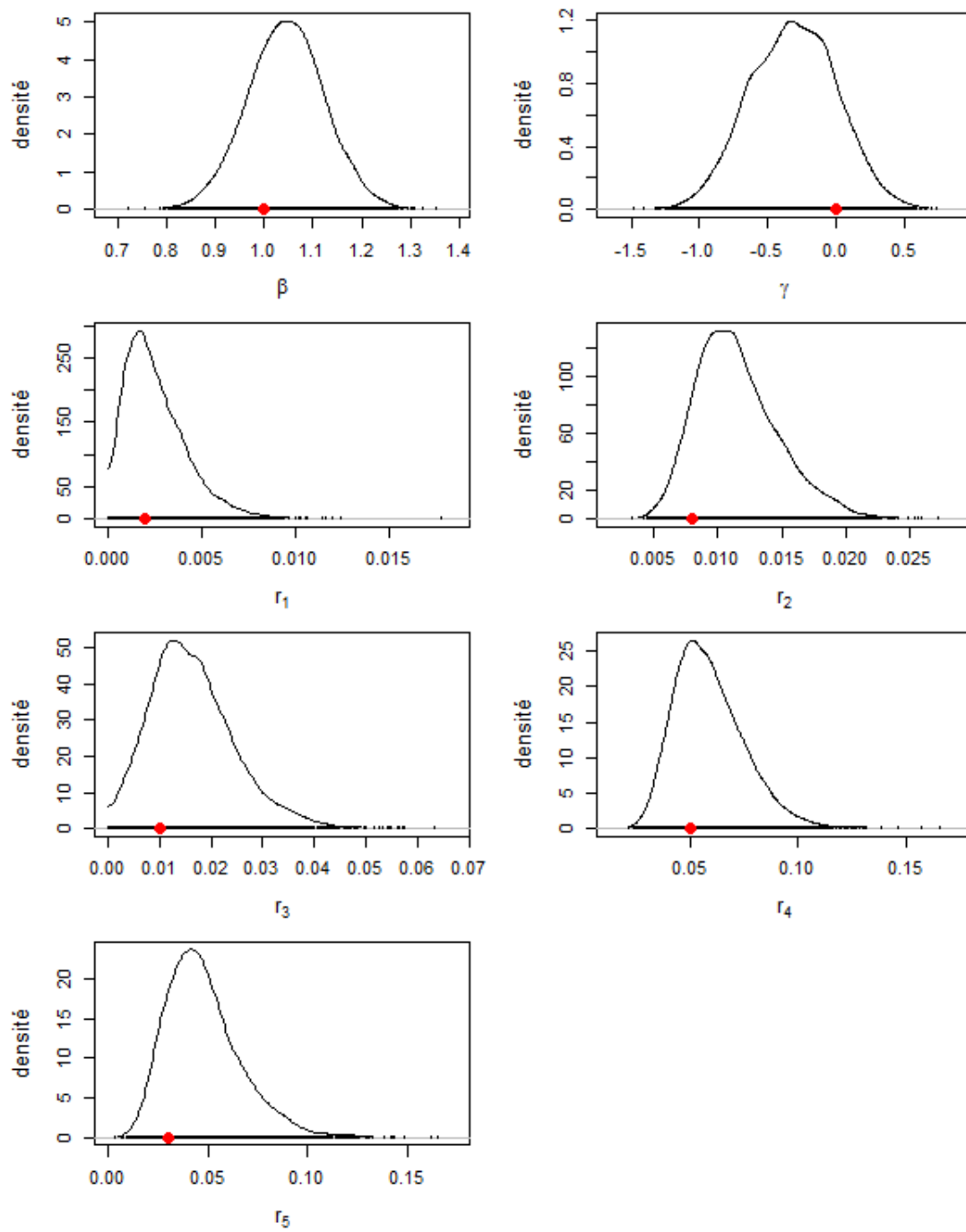


FIGURE 4.14 – Densité marginale a posteriori des paramètres sachant les données censurées simulées. Les vraies valeurs utilisées pour générer la simulation sont indiquées par des cercles rouges.

Le critère de Gelmann et Rubin à aussi été utilisé pour confirmer la convergence : pour toutes les chaînes $\hat{R} = 1$. Ici, les variances des lois de propositions pour chacun des paramètres de Θ sont : 0.1, 0.25, 0.005, 0.005, 0.03, 0.03, 0.05.

4.5 Analyse de sensibilité du modèle

Afin de répondre à l'objectif d'évaluation de l'influence des facteurs environnementaux et de la dispersion sur la dynamique d'invasion du pathogène nous avons réalisé des analyses de sensibilité globales grâce à de deux méthodes complémentaires suivant la méthodologie proposée par Sumner et al. [2012] : une première méthode de criblage qui permet d'explorer rapidement le comportement des sorties en faisant varier les paramètres d'entrée et une seconde méthode, plus coûteuse en temps de calcul, qui permet d'estimer les indices de Sobol i.e. estimer le pourcentage de variance expliquée par les paramètres sur les sorties d'intérêt. Les méthodes utilisées sont décrites en détail dans le Chapitre 1.

Nous avons choisi pour la première analyse la méthode de Morris [Morris, 1991] qui permet de calculer la moyenne μ^* des valeurs absolues des effets élémentaires de chaque entrée (qui permet de caractériser les paramètres ayant des effets importants en moyenne) et σ la dispersion des effets qui permet de caractériser les interactions et les effets non linéaires. Pour la seconde analyse nous avons choisi la méthode extended FAST [Saltelli et al., 1999] considérée comme la plus efficace en terme de nombre d'évaluations du modèle [Varella et al., 2010] [Makowski et al., 2006b]. Cette méthode permet d'évaluer d'une part les indices principaux de chaque paramètre i notés S_i et d'autre part les indices totaux notés ST_i qui permettent d'exprimer tous les effets de la variable d'entrée i sur les sorties d'intérêt.

Les sorties d'intérêt considérées dans ces analyses sont des nombres de patchs infectés (sorties scalaires) à différents temps et dans différentes zones spatiales ce qui permettra d'établir l'influence des paramètres temporellement et spatialement. Les indices sont calculés à trois dates (début, milieu et fin d'invasion) : 3 semaines (T3), 41 semaines (T41) et 83 semaines (T83) après la première semaine de détection, et sur trois zones géographiques : la Martinique entière (1236 patchs), la moitié Nord de l'île (618 patchs correspondant aux stations pluviométriques numéro 1 à 11 Figure 4.5) et la moitié Sud de l'île (618 patchs correspondant aux stations pluviométriques numéro 12 à 19).

Le modèle étant stochastique, nous avons répété 10 fois ces simulations afin de prendre en compte la variabilité des réalisations du modèle. Le calcul des indices a été fait sur les sorties d'intérêt moyennées.

Les paramètres (entrée du modèle) ont été étudiés dans les intervalles suivants : $\beta \in [0, 5]$, $\gamma \in [-5, 5]$, $r_1 \in [0, 10]$, $r_2 \in [0, 10]$, $r_3 \in [0, 10]$, $r_4 \in [0, 10]$, $r_5 \in [0, 10]$

4.5.1 Méthode de Morris

A l'aide du logiciel R [R Development Core Team, 2008] et du package "sensitivity" (Gilles Pujol), nous avons réalisé 100 tirages d'un plan d'expérience "One-At-a-Time" (trajectoires de Morris) sur des paramètres discrétisés en 4 niveaux. Calculer les indices de Morris pour les 7 paramètres de ce modèle a donc nécessité 800 simulations du modèle répétées 10 fois.

Cette phase de criblage permet d'identifier les paramètres les moins influents. Parmi les 7 paramètres, toutes zones géographiques et tous temps considérés confondus, les paramètres r_3 (réceptivité des jardins créoles), r_4 (réceptivité des plantations de Cavendish) et r_5 (réceptivité de plantains) ont de faibles effets sur la variabilité du nombre de patchs contaminés. Le paramètre β semble avoir systématiquement des effets non linéaires et/ou avec interaction (Figure 4.15).

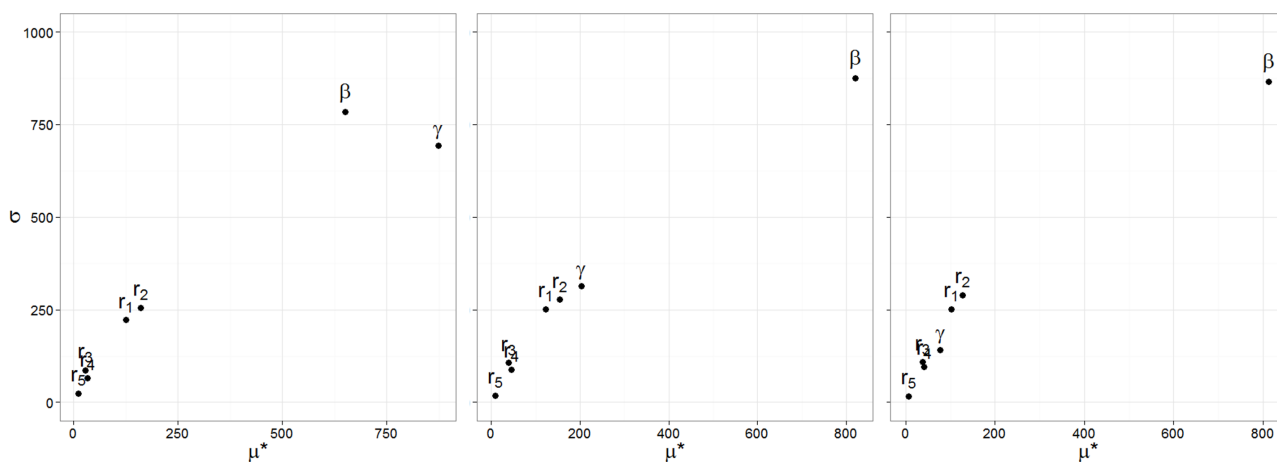


FIGURE 4.15 – Évaluation des indices μ^* et σ pour les 7 paramètres du modèle, calculés par la méthode de Morris sur le nombre de patchs infectés dans la zone Martinique au cours du temps : T3, T41, T83 (figures de gauche à droite), correspondant au début, au milieu et à la fin de l'invasion.

	Martinique			Zone Nord			Zone Sud		
	T3	T41	T83	T3	T41	T83	T3	T41	T83
β	2	1	1	2	1	1	2	1	1
γ	1	2	4	1	2	3	1	4	6
r_1	4	4	3	4	4	4	4	3	3
r_2	3	3	2	3	3	2	3	2	2
r_3	6	6	6	6	6	6	5	5	4
r_4	5	5	5	5	5	5	6	6	5
r_5	7	7	7	7	7	7	7	7	7

TABEAU 4.4 – Rang des effets de chacun des 7 paramètres selon les valeurs de μ^* . Le rang 1 (resp. 7) correspond aux plus fortes (resp. faibles) valeurs de μ^* .

Les influences relatives des 7 paramètres sont stables dans le temps à l'exception du paramètre γ , le plus influent en début de période mais qui n'intervient qu'en sixième position en fin de période en zone Sud et du paramètre β qui au contraire semble prendre de l'importance au cours de l'invasion (Tableau 4.4). Ceci peut être expliqué par l'accumulation des patches infectés au cours du temps : l'intensité du processus augmente avec le nombre de patches infectés (Équation 4.1). Il semble aussi y avoir un effet de la distribution spatiale : les trois paramètres peu influents (r_3 , r_4 , r_5) sont les paramètres relatifs aux types de production les moins présents dans les trois zones géographiques (Figure 4.6).

4.5.2 Méthode extended FAST

Une analyse par méthode e-FAST est proposée afin de confirmer les résultats précédents et d'évaluer plus précisément l'effet des paramètres sur le nombre de patches infectés au cours du temps. A l'aide du logiciel R et du package "sensitivity", chaque indice a été calculé par 500 simulations pour chaque paramètre (les valeurs des paramètres ont été tirés dans des lois uniformes sur les intervalles définis précédemment). Calculer les indices S_i et ST_i pour les 7 paramètres de ce modèle a donc nécessité 3500 simulations du modèle répétées 10 fois.

Les sorties (Tableau 4.6) confirment les premiers résultats obtenus avec la méthode de Morris : en se basant sur les indices ST_i (ou sur les rangs calculés Tableau 4.5), les paramètres les plus influents tous temps et toutes zones confondus sont les paramètres β et γ . L'effet temporel concernant ces deux paramètres semble être le même que celui décrit précédemment i.e. le paramètre γ lié à la pluie perd de l'importance au cours de l'invasion et inversement pour le paramètre β de la fonction de dispersion. Les paramètres liés au type de couverture du paysage semblent quant à eux prendre de l'importance au cours du temps.

	Martinique			Zone Nord			Zone Sud		
	T3	T41	T83	T3	T41	T83	T3	T41	T83
β	2	1	1	2	1	1	2	1	1
γ	1	3	2	1	2	2	1	2	2
r_1	3	3	5	3	4	5	4	3	5
r_2	4	4	6	4	3	6	3	4	3
r_3	6	7	7	7	7	7	6	7	7
r_4	5	6	4	5	5	4	5	6	6
r_5	7	5	3	6	6	3	7	5	4

TABLEAU 4.5 – Rang des effets de chacun des 7 paramètres du modèle selon les valeurs de ST_i . Le rang 1 (resp. 7) correspond aux plus fortes (resp. faibles) valeurs de ST_i .

A l'échelle de la Martinique, en début et en milieu d'invasion, les paramètres relatifs aux types de productions les moins fréquents sur le territoire (r_3 , r_4 , r_5) semblent être les moins influents (indices $ST_i < 0.006$ en T3 et $ST_i < 0.165$ en T41). En début d'invasion, le paramètre lié à la pluie est le plus influent ($ST_i > 0.91$) suivi par le paramètre β de la fonction de dispersion ($ST_i > 0.1$). En milieu d'invasion, les tendances s'inversent, le paramètre β devient le plus influent ($ST_i = 0.8$) suivi paramètre γ ($ST_i = 0.4$). En fin d'invasion, les tendances pour les paramètre β et γ restent les mêmes et tous les paramètres liés au type de production ont des valeurs d'indices $ST_i > 0.23$ (Tableau 4.6 et Figure 4.16).

Ces résultats suggèrent qu'en début et milieu d'épidémie, les deux types de production bananier spontané (r_1) et jardin privé (r_2) présents dans les patches influent sur le nombre de nouvelles infections (ST_i resp. > 0.03 à T3 et > 0.19 à T41). En fin d'invasion, tous les types de production sont influents ce qui suggère que l'influence des patches non colonisés va fortement dépendre du type de paysage qui les compose.

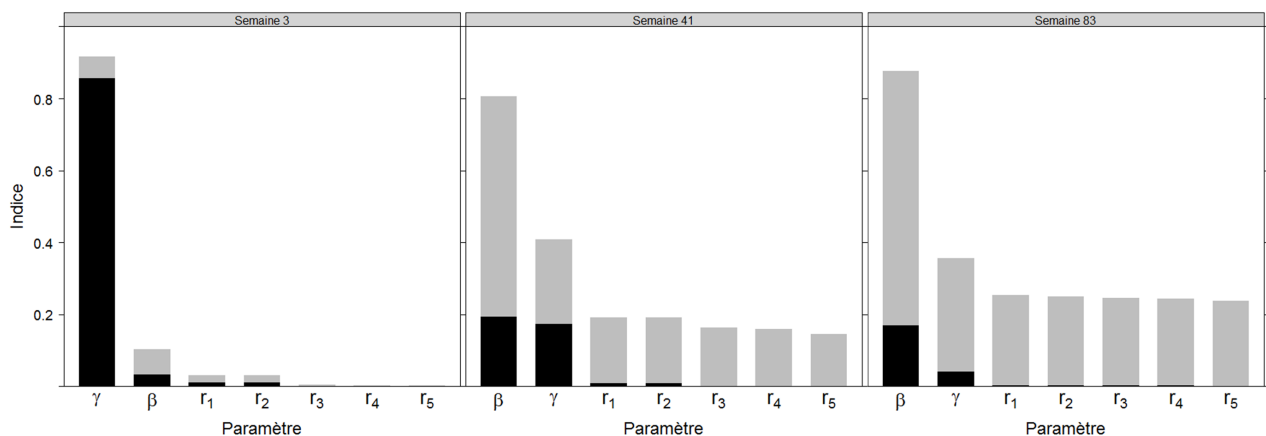


FIGURE 4.16 – Évaluation des indices S_i (en noir) et des interactions ($ST_i - S_i$) (en gris) pour les 7 paramètres du modèle, calculés par la méthode de Morris sur le nombre de patches infectés en zone Martinique au cours du temps : T3, T41, T83 (figures de gauche à droite), correspondant au début, milieu et fin d'invasion.

A l'échelle du Nord et du Sud de l'île, les mêmes observations sur le comportement des indices peuvent être données. On peut noter que les indices totaux des paramètres relatifs aux types de production semblent plus important en milieu d'invasion (T41) dans la zone Sud : $ST_i > 0.21$ que dans la zone Nord $ST_i > 0.1$. Ceci peut s'expliquer par la couverture du paysage différente dans les deux zones (cf. partie 4.2.3).

Pour conclure, les résultats de ces analyses e-Fast (à l'échelle de la Martinique découpée en deux zones ou non) montrent que l'influence des paramètres varie selon la période d'invasion. En effet, les résultats suggèrent une influence prépondérante de la pluie en début d'invasion mais aussi du paramètre de la fonction de dispersion et de deux types de production (jardins privés et bananiers spontanés). En milieu d'invasion, la dispersion (et toujours la pluie) prend de l'influence et en fin d'invasion, les 5 types de couverture deviennent influents par le biais d'interactions.

En effet, les paramètres liés au type de paysage ont des indices principaux S_i très faibles. La différence entre les valeurs des effets totaux ST_i et des effets principaux S_i prouve l'existence de fortes interactions entre les paramètres (Figure 4.16). Ces interactions peuvent s'expliquer par l'écriture mathématique du modèle, en effet la réceptivité d'un patch à la maladie est décrite par deux termes multiplicatifs (Équation 4.3.1).

Les deux méthodes d'analyse de sensibilité proposent des résultats quasi-similaires, les différences observées en comparant les rangs d'influence de chaque paramètre (Tableau 4.4 et Tableau 4.5) peuvent être expliquées par l'exploration limitée des entrées dans la méthode de Morris et la stochasticité du modèle.

	Martinique					
	T3		T41		T83	
	S_i	ST_i	S_i	ST_i	S_i	ST_i
β	0.0327	0.1026	0.194	0.80	0.16954	0.877
γ	0.8554	0.917	0.173	0.40	0.041	0.356
r_1	0.0106	0.0317	0.0085	0.191	0.0033	0.245
r_2	0.0104	0.0314	0.0081	0.19	0.0031	0.244
r_3	0.0002	0.0021	0.0003	0.145	0.0013	0.236
r_4	0.0013	0.0052	0.001	0.159	0.0018	0.249
r_5	0.0001	0.0019	0.0002	0.162	0.0017	0.253

	Zone Nord						Zone Sud					
	T3		T41		T83		T3		T41		T83	
	S_i	ST_i	S_i	ST_i	S_i	ST_i	S_i	ST_i	S_i	ST_i	S_i	ST_i
β	0.028	0.09	0.192	0.715	0.184	0.86408	0.037	0.117	0.175	0.88	0.14342	0.88
γ	0.87	0.93	0.301	0.565	0.084355	0.43398	0.83	0.897	0.0388	0.26	0.0104	0.32
r_1	0.006	0.0218	0.0127	0.1724	0.00453	0.23059	0.0167	0.047	0.004	0.218	0.00284	0.2713
r_2	0.005	0.0208	0.0125	0.1727	0.00447	0.22933	0.0166	0.0476	0.003	0.216	0.0027	0.2719
r_3	0.0002	0.0022	0.0006	0.1064	0.00078	0.21301	0.0003	0.0024	0.00042	0.191	0.002229	0.2692
r_4	0.0013	0.0063	0.0021	0.122	0.00176	0.23265	0.0012	0.0048	0.000568	0.2004	0.002222	0.2697
r_5	0.00016	0.0023	7.59e-05	0.1198	0.00139	0.23763	8.29e-05	0.0018	0.000562	0.2026	0.00227	0.2714

TABEAU 4.6 – Indices des effets principaux S_i et effets totaux ST_i pour chacun des 7 paramètres i du modèle dans les trois zones géographiques et sur trois temps d'invasion.

4.6 Inférence sur données réelles : Article en préparation

Après une validation robuste de l'algorithme 5 permettant l'inférence de données d'incidence censurées dans le temps et dans l'espace, nous mettons en application la méthode sur les données réelles (cf. section 4.2.1) et proposons en application la reconstitution de l'invasion de la cercosporiose noire en Martinique.

Ce travail est présenté sous forme d'un article "**A spatio-temporal stochastic model for an emerging plant disease spread in a heterogeneous landscape**" en cours d'écriture. En effet, les temps de calculs étant relativement longs, nous ne sommes pas en mesure de présenter à cette date les résultats finaux.

Nous projetons de le soumettre cet article à la revue *Bulletin of Mathematical Biology* dans les plus brefs délais.

A spatio-temporal stochastic model for an emerging plant disease spread in a heterogeneous landscape

Clara Landry · Catherine Abadie ·
François Bonnot · Jean Vaillant

Received: date / Accepted: date

Abstract Black Leaf Streak Disease (BLSD) or Black Sigatoka is the most severe foliar disease of banana and plantain. First detected in Fiji in 1963, this disease had a rapid worldwide expansion. In the Caribbean, BLSD has been still invaded new territories and it was first detected in the French lesser islands in 2010. BLSD is caused by *Mycosphaerella fijiensis*, an ascomycete fungus which produces wind-borne spores responsible for its spatial dispersal. In order to evaluate the BLSD long-distance dispersal and to better understand the effect of environmental factors on its spatial spread, a spatiotemporal study was set up during the recent BLSD invasion in Martinique. Disease observations were carried out in Martinique from September 2010 to May 2012 and sampling patches were defined from a regular spatial grid build over the island. In this paper, we consider a stochastic model of spatio-temporal propagation of BLSD in a heterogenous lanscape and we present mathematical and computational results for this continuous-time model. Statistical inference about parameters is carried out from presence-absence data using a bayesian framework based on an original data augmentation method with respect to patch first colonization times. A sensitivity analysis shows that the land cover pa-

Clara Landry
CIRAD Neufchateau
Tel.: +590-590-678910
Fax: +590-590-678910
E-mail: clara.landry@cirad.fr

Catherine Abadie
CIRAD Neufchateau
E-mail: catherine.abadie@cirad.fr

François Bonnot
CIRAD Montpellier
E-mail: francois.bonnot@cirad.fr

Jean Vaillant
Université des Antilles
E-mail: jean.vaillant@univ-ag.fr

rameters are less influential than the rainfall and dispersal ones on the BLSD invasion dynamics.

Keywords Stochastic process · Biological invasion · Likelihood · MCMC · Bayesian inference · Data augmentation

1 Introduction

Understanding biological pathogen invasion is a major contemporary issue: the appearance of new diseases cause serious disruption of existing ecosystems, biodiversity and local economy (Fisher et al, 2012) (Anderson et al, 2004). The evaluation of the spatiotemporal spread of a disease is necessary for effective and successful control strategies and requires a good knowledge on the long distance spread of the pathogen and on the evaluation of effects of environmental factors on this dispersion (Brown and Hovmoller (2002), Desprez-Loustau et al (2007)).

Most often, methods for fitting dynamic models have to take into account missing and censored data. However such missing data can be included as extra parameters in a bayesian framework. McKinley et al (2014) discussed about the use of data-augmented Markov chain Monte Carlo and presented a simulation-based bayesian inference for epidemic models illustrated using data from an outbreak of smallpox. Mollentze et al (2014) addressed some shortcomings of partially observed diseases by means of a reconstruction framework. Morelli et al (2012) developed a scheme for robust inference about transmission chains in the context of the Foot-and-Mouth Disease Virus. The framework presented leads to a Bayesian inference scheme able to reconstruct most likely transmission patterns and infection dates. Hoeting et al (2000) presented an auto logistic model with covariates for spatially correlated binary observations but the time aspect was not taken into account. They adopted a Bayesian set-up with incomplete observations from sampled responses taken over the area of interest. Their aim was to improve predictions of probability of presence for such lattice binary data. Walker et al (2006) demonstrated how methods of Markov chain Monte Carlo (MCMC) based on a hybrid sampler combining Gibbs sampling within Metropolis-Hastings frameworks can be used to derive estimates of parameters and missing data for ecological process models. Catterall et al (2012) used a continuous time discrete state space Markov process with spatially varying covariates to model colonization-dispersal of *Heracleum mantegazzianum*. They developed a bayesian approach to parameter inference taking into account uncertainty in colonization times.

In this paper, our objective is twofold. We first focus on modeling the invasion process of one of the major plant diseases affecting bananas and plantains, the Black Leaf Streak Disease (BLSD) or Black Sigatoka caused by the ascomycete fungus *Mycosphaerella fijiensis*, in an heterogeneous landscape island scale. Then, we tackle the problem of performing statistical inference on this model with incomplete data. In section 2, we present the presence-absence data available at the island scale over 83 survey weeks, and we introduce a

spatiotemporal model based on a stochastic intensity taking into account past neighboring infections and environmental variables. In section 3, the data augmentation technique we developed within a bayesian framework is presented and applied to the BLSD survey data in section 4. We conclude with a discussion in section 5.

2 Data and model

The Black Leaf Streak Disease (BLSD) or Black Sigatoka of bananas, due to the devastating ascomycete fungal plant pathogen *Mycosphaerella fijiensis*, is one of the most important food security threat at global scale (Pennisi, 2010). BLSD is specific to Musa hosts (bananas and plantain). The disease causes many lesions on leaves which become large necrotic surface reducing the photosynthetic surface of the plant. On susceptible cultivars, BLSD may destroy (by necrosis) respectively 40% to 100% of total leaf surface (Marin et al, 2003) which causes important yield losses, from 20 to 80% in the absence of fungicide treatments (Churchill, 2010).

Mycosphaerella fijiensis displays both an asexual and sexual reproduction emitting respectively and successively conidia and ascospores. Both spores play an important part in the fungus dispersal, but only ascospores are considered as responsible for the long-distance dispersal (Jones, 2000). It was recently demonstrated that the average dispersal is respectively few meters for conidia and 230 meters for ascospores but these spores can disperse up to 1 km. (Rieux et al, 2014)

BLSD is a recent and rapid global emergence, the disease was first described in 1963 in the Sigatoka Valley of Viti Levu, in the Republic of Fiji. A recent study proposes that Papua New Guinea would be the center of origin of BLSD (Robert et al, 2012). The disease spread rapidly on others continents as it was first described in Central America (Honduras) in 1972 and in Africa (Zambia) in 1973 since to most banana-producing regions worldwide. It is assumed that the very-long distance (continental) dispersal is due to human transportation of infected plant material.

In the Caribbean Islands, the first detection was in Cuba in 1990 and one of the most recent introduction was in Martinique (French west indies). The first report of BSLD on Martinique Island in September 2010 indicates that the disease was introduced passively by spores, from nearby islands where BLSD was yet present, during August 2010 (Ioos et al, 2011). Indeed, a surveillance network was set up since 2009 and allowed to check monthly the banana production against BLSD.

2.1 Data

Presence-absence data for *Mycosphaerella fijiensis* were collected in Martinique island within the framework of a disease surveillance network on ba-

nana over the whole island coordinated by the DAAF (Direction de l’Alimentation, de l’Agriculture et de la Forêt) and the FREDON Martinique (Federation REgionale de Defense contre les Organismes Nuisibles de la Martinique). Martinique island surface area is about 1228 km². A regular grid with cell size of 1 km x 1 km was placed over Martinique providing 1236 observation units. In the sequel, we refer to such an observation unit as a patch. Presence-absence observations were carried out in most of these patches (Fig. 1) on different plants and within different time windows (time between two observations of the same patch) and surrounding environmental variables were measured. After the detection of BLS on banana tree in a patch, no more observations were made on it because it is assumed that once a patch is colonized by the pathogen it remains infected. Five land-cover categories were defined : isolated banana from partially resistant varieties, private garden with few plants of partially resistant varieties, creole garden (small-scale farming with banana and other productions such as vegetables, fruits), plantain plots possibly extending over several hectares and Cavendish plots specialized in production for export on a larger scale. These plots are very susceptible to BLS and disease control are used to maintain good production. On Cavendish plots fungicide chemical treatment are preferably used whereas on plantain plots manual control, such as defoliating necrotic leaves, are privileged. The disease evaluation of each visual positive detection on a banana tree was confirmed with laboratory diagnostic. As lesions were analyzed on infected leaf samples using molecular markers of high sensitivity levels, there are no false-positive results.

Detection rates were spatially and temporally heterogeneous : the time window length was comprised between 1 and 82 weeks with an average of 21 weeks, in some patches only the date of the first detection was available and in other patches only the last date of non-detection was available. The whole observation period was over 83 weeks (from September 2010 to May 2012). Fig. 2 shows the spatial locations of the fungus at three different dates: beginning of the study with 37 patches contaminated, middle, and end of the study.

As environmental variables such as land-cover and humidity induced by rainfalls are assumed to affect the spreading process each of the 1236 patches covering Martinique was associated with a vector of covariables : the proportions of each landscape type and the rainfall level. The proportions of each landscape type were calculated from geo-referenced administrative surveys on Cavendish and plantain plots made from years 2010 to 2012 and expert opinions regarding the distribution of urban and rural zones of the island. We used also data, collected by Meteo France Martinique, on 19 weather stations located from north to south of the island. Daily rainfall quantities were averaged over week for each station. Each patch was assigned the rainfall level of the nearest station.

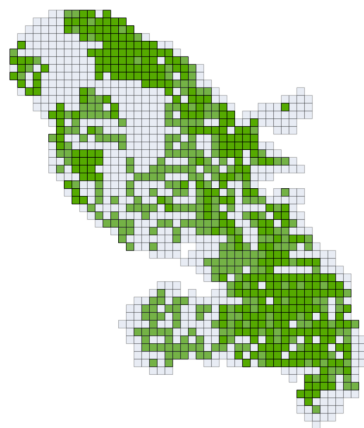


Fig. 1 Set of patches considered over the 83 weeks of survey. Green patches correspond to patches with positive or negative observations, grey patches are non-observed patches.

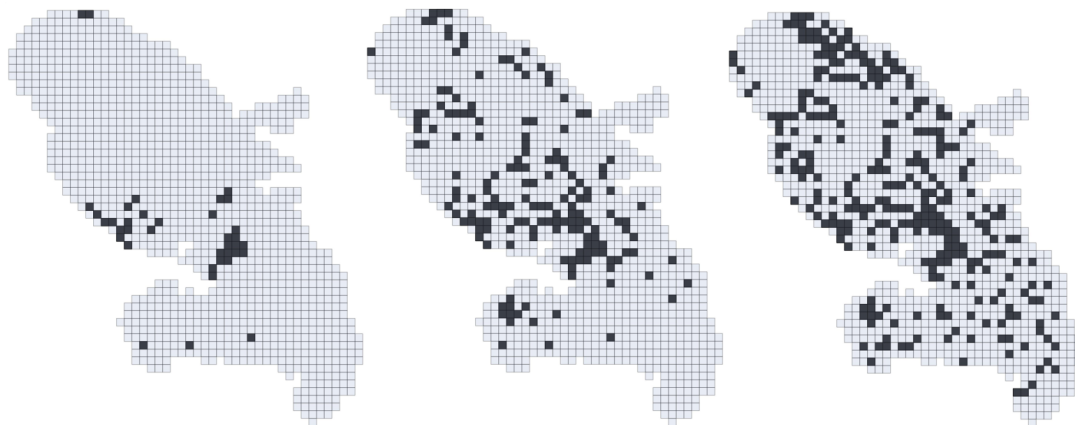


Fig. 2 Presence-absence maps for *Mycosphaerella fijiensis* at week 39 of year 2010, at week 29 of year 2011, and week 17 of year 2012 (left to right). Black squares are positive observation data and grey squares are uncolonizable, non informed or non infected data

2.2 Model and likelihood

We consider a spatiotemporal stochastic process of colonization which is assumed to be governed by its conditional intensity process. We denote by X the whole set of patches and by \mathcal{F}_{t-} the colonization history prior to time t so that the conditional intensity at any patch i of X and at time t in \mathbb{R}_+ conditional on \mathcal{F}_{t-} is defined as follows (e.g., Daley and Vere-Jones (2003)) :

$$\lambda(i, t | \mathcal{F}_{t-}) = \lim_{h \rightarrow 0^+} \frac{1}{h} Pr(\text{patch } i \text{ is colonized in } (t, t+h] \text{ given } \mathcal{F}_{t-}) \quad (1)$$

At time t , a susceptible patch i receives spores transmitted from patches that are already infected before t . We assume that this patch i has a suitability S_i associated with the rainfall level and the habitat receptivity to the infection. The effect of rainfall levels is modeled by an exponential term $\exp(\gamma \mathcal{P}_i)$ where \mathcal{P}_i is the mean rainfall level in patch i , and $\gamma \in (-\infty, \infty)$ is a parameter to be estimated : positive (negative) values of this parameter implies that locations with high rainfall are (not) preferred by the pathogen. The effect of habitat receptivity is modeled by a term $\sum_{k=0}^5 r_k \mathcal{H}_{i,k}$ where $\mathcal{H}_{i,k}$ is the proportion of patch i belonging to landscape category k , $k \in \{0, \dots, 5\}$ and $r_k \in [0, \infty)$ is a parameter representing the receptivity for colonization of landscape category k . We set $r_0 = 0$ to reflect that the volcano/sea (habitat 0) is uncolonizable due to the absence of host.

Let $d_{i,j}$ be the distance between the patches i and j . If j is infected before t and i susceptible at time t , the rate of transmission from patch j to patch i is assumed to be equal to $S_i f_\theta(d_{i,j})$ where $S_i = \exp(\gamma \mathcal{P}_i) \sum_{k=0}^5 r_k \mathcal{H}_{i,k}$ and $f_\beta(d_{i,j}) = \beta^2 d_{i,j}^{-2\beta}$ stands for the isotropic power-law dispersal kernel with decay parameter β .

Therefore, the stochastic intensity for a susceptible patch i at time t conditionally to the history of the process prior to time t is :

$$\lambda(i, t | \mathcal{F}_{t-}) = \exp(\gamma \mathcal{P}_i) \sum_{k=0}^5 (r_k \mathcal{H}_{i,k}) \sum_{j \in C_{t-}} f_\beta(d_{i,j}) \quad (2)$$

where C_{t-} is the set of colonized patches before time t .

It is worth noticing that if patch i is colonized at time t , then

$$\lambda(i, t | \mathcal{F}_{t-}) = 0.$$

Denoting by θ the parameter vector, we have $\theta = (\beta, \gamma, r_1, \dots, r_5)$. We face the problem of inferring θ from incomplete data consisting of intervals in which first dates of colonization are detected or not for patches included in the survey over the time interval $[0, T]$. If we denote by τ_i the unobserved colonization date for patch i , then the complete data set is

$$\tau = \{\min(\tau_i, T), i \in X\}$$

and the likelihood function of the complete complete data set is (see for example Cressie and Wikle (2011), p 349)

$$L(\theta; \tau) = \left(\prod_{\substack{i \in X \\ \tau_i \leq T}} \lambda(i, \tau_i | \mathcal{F}_{\tau_i-}) \right) \times \exp\left(- \sum_{i \in X} \int_0^T \lambda(i, t | \mathcal{F}_{t-}) dt\right) \quad (3)$$

From (2) and (3), we can write

$$L(\theta; \tau) = \prod_{i \in C_T} \left[\left(S_i \sum_{j \in C_{\tau_i-}} f_{\beta}(d_{i,j}) \right) \times \exp \left\{ -S_i \sum_{j \in C_{\tau_i-}} f_{\beta}(d_{i,j})(\tau_i - \tau_j) \right\} \right] \quad (4)$$

$$\times \prod_{i \notin C_T} \exp \left\{ -S_i \sum_{j \in C_{T-}} f_{\beta}(d_{i,j})(T - \tau_j) \right\}$$

where $S_i = \exp(\gamma \mathcal{P}_i) \sum_{k=0}^5 (r_k \mathcal{H}_{i,k})$ is the suitability for colonization in patch i .

The first product on the right-hand side of (4) is the contribution from the patches colonized before the last observation date T whereas the second one corresponds to the probability for the remaining patches not being colonized at this date.

We have to tackle the problem of inferring θ from incomplete data consisting of intervals in which dates of colonization occur. The complete data set τ is not available and the incomplete data likelihood involves integrals which are analytically and numerically intractable (6). If we denote by \mathcal{D} the observed incomplete set of data, then the joint posterior distribution for θ and τ is

$$P(\theta, \tau | \mathcal{D}) = \frac{P(\mathcal{D}, \tau | \theta) \pi(\theta)}{P(\mathcal{D})} \quad (5)$$

where $\pi(\cdot)$ is the prior distribution for θ and $P(\mathcal{D})$ considered as a normalizing constant. $\pi(\cdot)$ reflects the knowledge about the model parameters prior to the observations. Therefore, the posterior distribution for θ is obtained by marginalizing over τ :

$$P(\theta | \mathcal{D}) = \int_H P(\theta, \tau | \mathcal{D}) d\tau. \quad (6)$$

where H is the space of unobserved colonization times.

In next section, we focus on MCMC techniques (Robert and Casella, 2010) for sampling from the posterior $P(\theta, \tau | \mathcal{D})$ and then from $P(\theta, | \mathcal{D})$. For any single parameter, the marginal posterior is then estimated by the sampled values of this parameter only.

3 MCMC

By using MCMC techniques and considering the unknown colonization times as augmented variables in the MCMC algorithm, we can sample from the posterior distribution of θ . We generalized the sampling method proposed by Catterall et al (2012) by updating the whole set of first colonization times at each iteration by means of a reversible-jump hybrid sampler. This updating method reconstructs realizations of the colonization times consistent with the

Table 1 Notation used in the likelihood and the MCMC algorithm

Notation	Definition
τ_i	time of colonization of the i^{th} patch
τ_{end}	the time the last observation was made, in week
C_t	the set of patches colonized up to time t
C_{t-}	the set of patches colonized prior to time t
N	the number of patches in Martinique
$\lambda(i, t \mathcal{F}_{t-})$	the rate of colonization of a patch i in time t conditionally to the history of the process prior to time t \mathcal{F}_{t-} . (denoted by $\lambda(i, t)$ for simplicity)
θ	$= (\beta, \gamma, r_1, \dots, r_5)$, the parameter vector
H	the space of unobserved colonization times
$\pi(\cdot)$	prior distribution of parameter
$q_k(\cdot, \cdot)$	proposal distribution for the k th element of θ

data and the model. Each iteration of this algorithm includes also updating of all model parameters. We assume that the prior distributions for the parameters are independent and uniform over a fixed interval. To sample from the posterior distribution $P(\theta, \tau | \mathcal{D})$, the algorithm starts with an initial vector $\tau^{(0)}$ of colonization times consistent with the observation \mathcal{D} . Then, at the m th step, the vector $\tau^{(m+1)}$ of colonization times consistent with \mathcal{D} is obtained from a sampler conditional on $\tau^{(m)}$ and $\theta^{(m)}$ the parameter vector sampled at step m . At step m , our proposal samples a vector $\tau^{(m+1)}$ according to a truncated multivariate Gaussian law centered on $\tau^{(m)}$ with an adaptive variance matrix. The algorithm is as follows:

1. Start with $\tau^{(0)}$ consistent with the data and iterate the following procedure:
2. Assign values to $\theta^{(0)}$;
3. Set $m = 0$;
4. Repeat
 - Draw $\tau^{(m+1)}$ consistent with \mathcal{D} using the data augmentation sampler with $\tau^{(m)}$ and $\theta^{(m)}$;
 - Sample $\theta^{(m+1)}$ using Metropolis-Hastings with $\tau^{(m+1)}$ and $\theta^{(m)}$;
 - Set $m = m + 1$;
 - Store every $\tau^{(m)}$ and $\theta^{(m)}$; after an initial burn-in period;
5. End repeat.

For $\tau^{(0)}$, each unknown colonization time is drawn according to the uniform distribution on a interval consistent with \mathcal{D} .

$\theta^{(m+1)} = (\beta^{(m+1)}, \gamma^{(m+1)}, r_1^{(m+1)}, r_2^{(m+1)}, r_3^{(m+1)}, r_4^{(m+1)}, r_5^{(m+1)})$ and the k th element $\theta_k^{(m+1)}$ of $\theta^{(m+1)}$ is drawn according to the acceptance probability

$$\min\left(1, \frac{P((\theta_1^{(m+1)}, \dots, \theta_{k-1}^{(m+1)}, \theta_k^*, \theta_{k+1}^{(m)}, \dots, \theta_7^{(m)}); \tau^{(m+1)}) q_k(\theta_k^*, \theta_k^{(m)})}{P((\theta_1^{(m+1)}, \dots, \theta_{k-1}^{(m+1)}, \theta_k^{(m)}, \theta_{k+1}^{(m)}, \dots, \theta_7^{(m)}); \tau^{(m+1)}) q_k(\theta_k^{(m)}, \theta_k^*)}\right) \quad (7)$$

where q_k is the proposal for θ_k .

4 Results

The prior distributions chosen for the model parameters reflect the lack of information about these parameters : $\beta \sim U[0, 10]$, $\gamma \sim U[-10, 10]$ and $r_k \sim U[0, 10]$ for each k in $\{1, \dots, 7\}$.

We used Normal distributions centered on the current parameter for each proposal distribution. After tuning, the standard deviation used for $\beta, \gamma, r_1, r_2, r_3, r_4$ and r_5 are respectively 0.01, 0.5, 0.001, 0.001, 0.01, 0.01 and 0.01. Fig. 3 give the posterior distribution of the land cover parameters, and Fig. 4 the ones of β and γ . The posterior means and standard errors are given in table 2. From these parameter estimates, numerical simulations were carried out. Fig. 5 shows three examples.

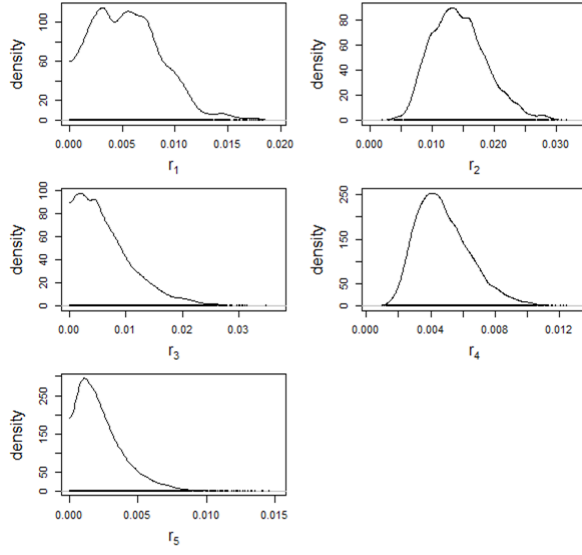


Fig. 3 Bayesian distribution of the land cover parameters.

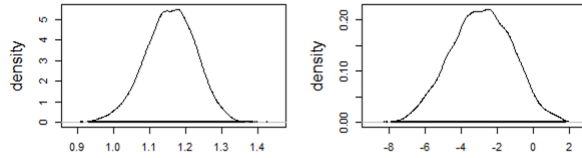


Fig. 4 Posterior distributions : of β (on the left); of γ (on the right).

New infected patches tend to be clustered around previous ones probably because of the pathogen spread characteristics. The invasion process seems to

be related to the landscape structure : the southern part of the island is less infected by the disease, probably because of the dry vegetation of this region whereas the northern part is mostly covered by the tropical rain-forest.

It is worth noticing that the model can be easily simulated as a time continuous Markov process when the covariates vary spatially only. We use the fact that the waiting time between two consecutive events follows conditionally to the past history an exponential distribution with parameter depending on the set of patches colonized prior to the last event. Fig. 5 shows artificial data generated from the model with bayesian parameter estimates.

Table 2 Bayesian estimates of θ

Parameter	Posterior Mean	Posterior standard error
β	1.158	0.070
γ	-2.853	1.688
r_1	5.520e-03	3.268e-03
r_2	1.440e-02	4.437e-03
r_3	6.542e-03	5.101e-03
r_4	4.793e-03	1.667e-03
r_5	2.310e-03	1.742e-03

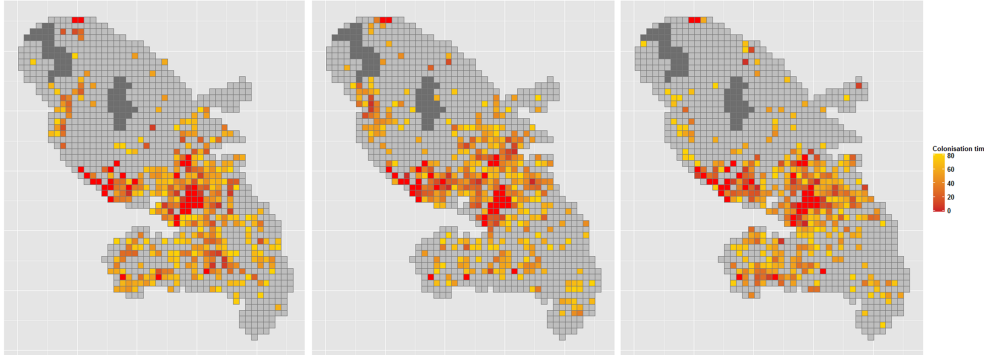


Fig. 5 Three simulated presence-absence maps over 83 weeks from the model with our bayesian estimated parameters. The color indicates the first colonization period in each patch

To better understand the contribution of each parameter and each interaction between parameters to the variance of the number of infected patches in three times, we carried out a global sensitivity analysis (GSA) using variance-based method: the extended Fourier Amplitude Sensitivity Test (eFAST) (Saltelli et al, 1999). Methodology for GSA is well described in (Saltelli et al, 2004). In this method, all the parameters (7) were varied simultaneously in the vari-

ation space given as follow: $\beta \in [0, 5]$, $\gamma \in [-5, 5]$, $r_1 \in [0, 10]$, $r_2 \in [0, 10]$, $r_3 \in [0, 10]$, $r_4 \in [0, 10]$, $r_5 \in [0, 10]$. For each output, two indexes are calculated. The first-order sensitivity index (or "main effect") measuring the contribution to the output variance of the parameter θ_i alone and the interaction index measuring the difference between the total effect (contribution to the output variance of θ_i including all variance caused by its interactions) and the first-order effect of the parameter. We analysed the model sensitivity indexes using 3500 simulations repeated 10 times to take account about stochasticity.

Parameters identified as the most influential were the parameter β of the dispersion function, the parameter γ relative to the rain value in the patch and the land-cover parameters r_1 and r_2 relative to the isolated banana and the private garden productions. The analysis of the sensitivity indexes within time suggest the increase in the influence of the cover-land parameters. The high interactions between parameters can be explained by the model receptivity description: the two covariates effects are taken to be multiplicative.

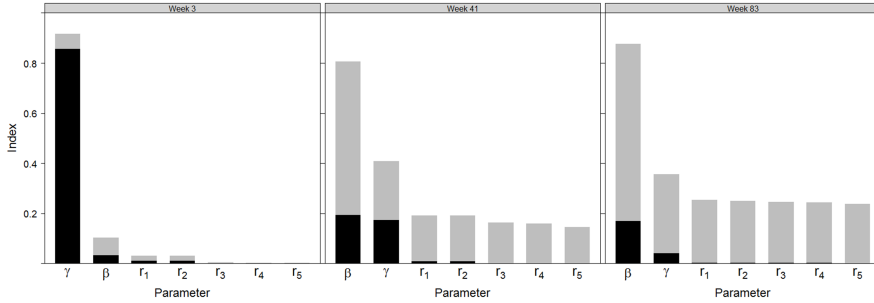


Fig. 6 Sensitivity indices of the eFAST method in three times: week 3, week 41 and week 83 corresponding to the beginning, the middle and the end of the invasion. In black, first-order effects of each parameter, in grey, interactions.

5 Discussions

The general objective of the paper was to model the invasion process of an emerging plant disease heterogeneous landscape. To reach this aim, we had to tackle the problem of performing statistical inference on the model with many incomplete data. These approach was applied on the incidence data obtained during the invasion of Black Leaf Streak Disease of bananas in Martinique.

We modeled successfully the BLSD invasion through a spatiotemporal stochastic process. In the framework of metapopulation model, the island landscape was divided in patches of 1km^2 which were characterised by two environmental factors (rainfall and habitat type).

In the case of BLSD invasion dynamics in Martinique, we carried out a sensitivity analysis which shows that the land cover are less influential than the

rainfall and dispersal parameters. The sensitivity analysis had to be performed on averaged outputs because of the stochasticity of the model. Marrel et al (2012) proposed the using of metamodels in this context to obtain valuable indexes.

We used the Bayesian paradigm for fitting a spatio-temporal stochastic process to patch data with unobserved first colonization times and subject to censoring. Bayesian statistical inference could be performed owing to a data augmentation method which reconstructs the first colonization time for each patch at each iteration of MCMC algorithm. This data reconstruction method is powerful and original in the epidemiology science. Whereas some short-distance function were available for BLS (Rieux et al, 2014), the inference of the dispersal parameter allowed to get for the first time a long-distance function for this disease. This function would be useful to develop spatial epidemiological model to evaluate control methods at landscape scale (such as varieties mixtures).

Numerical simulations can be performed to evaluate where future colonizations are likely to occur using the bayesian parameter estimates obtained from our inferential routine. We could simulate patch colonizations beyond the end of our survey data and for unobserved patches. These predictive data generation can be useful in a control strategy framework. The numerical simulation can be also carry out in case where the covariates are time dependent but this is more complicated and this research work is in progress.

References

- Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* 19(10):535 – 544
- Brown J, Hovmoller M (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* 297:537–541
- Catterall S, Cook AR, Marion G, Butler A, Hulme PE (2012) Accounting for uncertainty in colonisation times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. *Ecography* 35(10):901–911
- Churchill A (2010) *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular plant pathology* pp 1–22
- Cressie N, Wikle CK (2011) *Statistics for Spatio-Temporal Data*. Wiley
- Daley D, Vere-Jones D (2003) *An Introduction to the Theory of Point Processes*. Springer, 2nd Ed., New York
- Desprez-Loustau M, Robin C, Buee M, Courtecuisse R, Garbaye J, Suffert F, Sache I, DM R (2007) The fungal dimension of biological invasions. *Trends in Ecology & Evolution* 22(9):472–80

- Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL, Gurr SJ (2012) Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484(7393):186–194
- Hoeting JA, Leecaster M, Bowden D (2000) An improved model for spatially correlated binary responses. *Journal of Agricultural, Biological, and Environmental Statistics* 5(1):102–114
- Ioos R, Hubert J, Abadie C, Duféal D, Opdebeeck G, Iotti J (2011) First report of black sigatoka disease in banana caused by *mycosphaerella fijiensis* on Martinique island. *Plant Disease* 95(3):359–359
- Jones D (2000) Diseases of banana, abaca and enset, CABI publishing, chap 2, pp 37–142
- Marin DH, Romero RA, Guzman M, Sutton TB (2003) Black sigatoka: An increasing threat to banana cultivation. *Plant Disease* 87(3):208–222
- Marrel A, Ioos B, Da Veiga S, Ribatet M (2012) Global sensitivity analysis of stochastic computer models with joint metamodels. *Stat Comput* 22:833–847
- McKinley TJ, Ross JV, Deardon R, Cook AR (2014) Simulation-based bayesian inference for epidemic models. *Computational Statistics and Data Analysis* 71:434 – 447
- Mollentze N, Nel L, Townsend S, le Roux K, Hampson K, Haydon D, Soubeyrand S (2014) A bayesian approach for inferring the dynamics of partially observed endemic infectious diseases from space-time-genetic data. *Proceedings of the Royal Society of London Series B: Biological Sciences* 281(1782)
- Morelli MJ, Thébaud G, Chadoeuf J, King DP, Haydon DT, Soubeyrand S (2012) A bayesian inference framework to reconstruct transmission trees using epidemiological and genetic data. *PLoS Comput Biol* 8(11)
- Pennisi E (2010) Armed and dangerous. *Science* 327(5967):804–5
- Rieux A, Soubeyrand S, Bonnot F, Klein EK, Ngando JE, Mehl A, Ravigne V, Carl, de Lapeyre de Bellaire L (2014) Long-distance wind-dispersal of spores in a fungal plant pathogen: estimation of anisotropic dispersal kernels from an extensive field experiment. *Plos One*
- Robert C, Casella G (2010) Monte Carlo Statistical Methods. second ed. Springer
- Robert S, Ravigne V, Zapater MF, Abadie C, Carlier J (2012) Contrasting introduction scenarios among continents in the worldwide invasion of the banana fungal pathogen *mycosphaerella fijiensis*. *Molecular Ecology* pp 11–12
- Saltelli A, Tarantola S, Chan K (1999) A quantitative model independent method for global sensitivity analysis of model output. *Technometrics* 41:39–56
- Saltelli A, Tarantola S, Campolongo F, Ratto M (2004) Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models. WILEY
- Walker DM, Perez-Barberia FJ, Marion G (2006) Stochastic modelling of ecological processes using hybrid gibbs samplers. *Ecological Modelling* 198:40 – 52

4.7 Discussion du chapitre

L'objectif du présent chapitre était d'évaluer la dispersion à longue distance et de mieux comprendre l'effet du climat et de la couverture du paysage dans la dynamique d'invasion de la cercosporiose noire en Martinique. Pour cela et compte-tenu des données disponibles pour cette étude, un développement technique important a dû être réalisé.

En effet, dans ce chapitre nous avons montré comment des données d'incidence censurées recueillies à grande échelle spatiale peuvent permettre de mieux comprendre l'invasion d'un pathogène dans un paysage hétérogène. Nous avons utilisé un processus ponctuel spatio-temporel en temps continu qui nous a permis de simuler l'invasion sur un territoire discrétisé en patches, d'inférer un paramètre de dispersion et des paramètres environnementaux jugés a priori influents dans l'invasion, et de valider nos a priori en analysant l'influence de ces paramètres sur le nombre de patches infectés au cours du temps. Dans l'article en préparation présenté, une autre application du modèle et de son inférence est proposée : connaissant les intervalles de temps dans lesquels la colonisation des différents patches a eu lieu, l'objectif est la reconstruction de l'histoire de l'invasion du pathogène.

4.7.1 Discussion des principaux résultats

Informations acquises sur la dynamique d'invasion de la cercosporiose noire. Expérimentalement, il est très difficile, voir impossible, d'acquérir ces connaissances à l'échelle d'un territoire. Notre démarche de modélisation a permis de pallier le problème et d'acquérir des connaissances à l'échelle d'un territoire sur la dispersion du pathogène responsable de la cercosporiose noire et de l'effet de facteurs environnementaux : la pluviométrie et la couverture du paysage (type de productions bananières).

Fonction de dispersion à longue distance. Sur ce pathogène, des fonctions de dispersion à l'échelle d'un kilomètre ont déjà été décrites [Rieux et al., 2014]. Nous avons la connaissance que le pathogène peut se disperser à plus de 6 km [Amil et al., 2007]. Cependant, l'inférence du paramètre de la fonction de dispersion de type puissance (proposée dans le modèle développé) permet d'obtenir, pour la première fois pour cette maladie, une fonction de dispersion à longue distance.

Facteurs environnementaux. Les analyses de sensibilité du modèle développé montrent (de façon similaire pour la méthode de Morris et e-Fast) que, sachant les domaines de variations des paramètres proposés, les paramètres qui expliquent la majeure partie de la variance observée du nombre de patches contaminés sont les paramètres relatifs à la fonction de dispersion et celui relatif aux hauteurs pluviométriques.

Les résultats des analyses de sensibilité montrent également un effet des types de production spontané et jardin privé. L'inférence sur les données réelles révèlent que la moyenne a posteriori

la plus élevée, concerne le type de production jardin privé ($r_2 = 1.44e^{-2}$), autrement dit, ce type de production semble être celui ayant contribué le plus au succès d'invasion de la maladie en Martinique.

Il est intéressant de noter que ce type de production compte parmi les plus représentées sur le territoire. En effet, plus il y a des bananiers répartis spatialement, plus grande est la vitesse d'invasion. Ces résultats sont en accord avec l'expérience réussie d'éradication de la cercosporiose noire en Australie qui a consisté à éliminer tous les bananiers isolés et remplacer les bananiers traditionnels dans les jardins privés par des variétés résistantes [Henderson et al., 2006].

On peut également rapprocher ces résultats de ceux obtenus lors de l'analyse des données historiques de l'invasion de la cercosporiose noire dans les Caraïbes acquises par des enquêtes régionales. En considérant la superficie totale de chaque territoire et la durée d'invasion sur ce territoire, on peut estimer indirectement la vitesse d'invasion de la maladie en Martinique à 49 km²/mois et 88 km²/mois à Trinidad. Cette vitesse d'invasion deux fois plus lente en Martinique pourrait s'expliquer par des différences de couverture du paysage et/ou de conditions pluviométriques (Abadie, communication personnelle).

L'influence de la répartition spatiale de l'hôte sur la dispersion des maladies a également été montrée sur une autre maladie fongique à dispersion aérienne : sur une surface de 15 hectares, [Mundt et al., 2011] montrent que l'invasion de la rouille du blé peut être stoppée par une diminution de la densité d'hôtes.

Dans des travaux théoriques récents, Strona [2015] étudie par modélisation analytique l'effet du taux de colonisation et de la distance de dispersion sur la dynamique d'invasion. En fonction du taux de colonisation, les stratégies de dispersion de l'espèce peut être différente. Pour notre étude, nous ne disposons pas de données quantitatives de sévérité de la maladie qui auraient pu être prises en compte dans le modèle.

Informations acquises sur les réseaux de surveillance. Ce cas d'étude suggère une analyse a posteriori de la performance de détection de la maladie dans le réseau d'épidémiologie. Jusqu'à sa détection, la surveillance était basée sur un réseau de vingt parcelles fixes de plantains et des parcelles aléatoires de type Cavendish. Cependant, cinquante hectares ont été détectés positifs deux semaines après la première détection. Cela signifie que la maladie était présente avant, et qu'il aurait fallu concentrer la surveillance surtout sur les autres types de cultures bananières non commerciales dont on a montré l'influence. Parnell et al. [2012] et Parnell et al. [2013] proposent une méthode statistique appropriée pour élaborer des programmes de surveillance.

4.7.2 Discussion méthodologique

Les données. Nous avons utilisé dans ces analyses toutes les données disponibles. Les données d'incidences décrites (section 4.2.1) sont censurées dans l'espace puisque tous les patchs définis

n'ont pas été visités, et dans le temps puisque nous disposions seulement de fenêtres de temps dans lesquelles les colonisations ont eu lieu.

Une fois la maladie détectée, la confirmation par analyse moléculaire ne laisse aucun doute sur la réelle contamination (fiabilité des détections positives). On peut cependant émettre l'hypothèse d'un manque de fiabilité des observations négatives. En effet, bien que les lésions causées par *Mycosphaerella fijiensis* soient identifiables, il est toujours possible qu'un observateur ne détecte pas la maladie. [Parnell et al. \[2012\]](#) proposent dans le cas de maladie asymptomatique une méthode statistique d'échantillonnage.

Les données météorologiques de pluviométrie proviennent de 19 stations. Nous avons donc établi 19 zones en Martinique en attribuant à chaque patch la hauteur de pluie de la station la plus proche. Ces zones sont critiquables car elles ne tiennent pas compte de la réelle topologie de l'île. Les données de vent quant à elles n'ont pu être exploitées faute d'informations suffisamment précises issues des stations météorologiques mesurant le vent.

Mycosphaerella fijiensis étant un pathogène spécifique du bananier, nous avons décrit le paysage au moyen de types de production bananière en utilisant des informations d'expertises et le parcellaire déclaré auprès des autorités. Cependant les productions déclarées, notamment pour le type de production plantain, sont en quantités très inférieures aux quantités réelles. On ne peut donc pas négliger un biais dans la description paysagère de la Martinique.

Le modèle. Les modèles basés sur des processus ponctuels stochastiques offrent un cadre bien adapté pour décrire une invasion en utilisant des données spatio-temporelles d'incidence d'une maladie [[Daley and Vere-Jones, 2003](#)] [[Vaillant et al., 2010](#)] [[Cook et al., 2007](#)]. Dans le modèle développé, il est admis que le patch est infectieux dès lors qu'il est colonisé. Plus précisément, on suppose que les lésions d'une plante détectée positive sont sporulantes (et de surcroît en ascospores, spores sexuées terminant le cycle infectieux du pathogène), or ce n'est sans doute pas le cas. Il serait alors approprié d'insérer dans le modèle une période de latence sexuée entre la contamination d'un patch et son état infectieux, ce qui peut être réalisé en insérant un état Latent entre nos deux états Sain - Infectieux [[Höhle et al., 2005](#)].

La prise en compte de covariables environnementales et plus particulièrement l'hétérogénéité du paysage est un atout majeur dans la modélisation spatio-temporelle d'épidémie ([Fortin et al. \[2003\]](#), [Plantegenest et al. \[2007\]](#)). La pluviométrie a été introduite dans le modèle sous l'hypothèse d'un effet exponentiel sur la capacité d'un patch à recevoir des spores et l'hétérogénéité du paysage a aussi été introduite dans la réceptivité des patches en fonction du type de production qui les composent. Cette réceptivité prend donc aussi en compte les méthodes de lutte utilisées. Dans le modèle, ces données sont fixées à leur valeur moyenne sur les 83 semaines d'étude. Il serait intéressant de prendre en compte l'aspect temporel de ces covariables (les données hebdomadaires étant disponibles) qui permettrait d'affiner le processus.

Une autre amélioration possible du modèle serait de considérer des capacités d'émission de

spores dans chacun des patches. Ces capacités d'émission dépendraient de la proportion des types de production présents dans chaque patch et influeraient sur l'intensité du processus en ajoutant cinq paramètres supplémentaires. Une première amélioration du modèle a été réalisée en incluant les capacités d'émission des spores des patches contaminés. Les analyses de sensibilité (e-Fast) de ce modèle amélioré corroborent les résultats du modèle simple et n'apportent pas d'informations supplémentaires sur l'effet des couvertures du paysage. La prise en compte du caractère infectieux des patches devra être modélisée autrement.

L'utilisation d'une fonction de dispersion isotrope de type puissance permet de simuler une propagation à longue distance. Compte tenu des possibles modes de dispersion de la maladie, il serait aussi intéressant de considérer dans le modèle un inoculum extérieur (qui pourrait être constant [Brown et al., 2013]) et qui refléterait le fait que la maladie peut être introduite par des vents provenant des îles contaminées voisines. En effet, des analyses génétiques du pathogène en Martinique montrent que les populations invasives proviennent de Sainte Lucie, l'île à 50 km au sud de la Martinique. Cette invasion extérieure a pu se produire de façon régulière tout au long de la période d'étude.

Inférence des paramètres du modèle Les données étant incomplètes, nous avons utilisé une démarche "pas-à-pas" pour inférer les paramètres du modèle grâce à des algorithmes MCMC de Hastings-Métropolis. Une première étape a été d'implémenter un algorithme qui considère des données d'incidence où les temps de colonisation sont connus (décrit par Cook et al. [2007]). La seconde étape a été d'implémenter un algorithme qui permet de prendre en compte nos données censurées en intervalles (décrit par Catterall et al. [2012] et Brown et al. [2013]). Ces algorithmes ont été validés sur plusieurs jeux de données simulées : les modes des densités a posteriori concordent avec nos paramètres d'entrée connus.

L'inférence utilisant des données censurées à grande échelle spatiale et permettant d'obtenir des estimations robustes des paramètres est, à notre connaissance, très peu utilisée en épidémiologie végétale bien que l'incertitude des temps de colonisation soit inhérente aux observations biologiques. Cette étude peut-être considérée comme originale sur ce point.

Le problème majeur de cette méthode d'inférence concerne les temps de calcul, en effet, devoir proposer à chaque itération de l'algorithme un temps de colonisation dans chacune des fenêtres d'observation de chacun des patches augmente considérablement les temps de calcul. Brown et al. [2013] évaluent différentes techniques (par exemple : borner la fonction de dispersion ou bien utiliser un code parallélisé) pour pallier le problème. McKinley et al. [2014] présentent une méthode DA-MCMC (data-augmented Markov chain Monte Carlo) utilisant un échantillonnage préférentiel qui s'appliquerait à notre cas d'étude et qui permettrait l'optimisation de la convergence des chaînes de Markov et donc l'amélioration des temps de calculs.

Concernant l'inférence des paramètres sur les données réelles, les résultats sont en cours de réalisation, le nombre d'itérations est encore trop faible pour pouvoir conclure sur les densités a

posteriori des paramètres. Comme dans [Catterall et al. \[2012\]](#) ou [Meentemeyer et al. \[2011\]](#) ces densités vont d'une part, permettre d'évaluer le type de paysage ayant contribué le plus à l'invasion de la maladie (ce qui semble être les jardins privés) et d'autre part, permettre de proposer des scénarios de reconstruction de l'histoire de l'invasion de la cercosporiose noire en Martinique.

Analyse de sensibilité Dans une optique de meilleure compréhension du processus d'invasion décrit, nous avons analysé l'effet de nos paramètres sur la variabilité du nombre de patch infectés au cours des 83 semaines d'étude, en utilisant deux méthodes d'analyse de sensibilité. Le modèle étant stochastique, nous avons choisi de réaliser dix simulations pour chacune des combinaisons de paramètres et pour chaque méthode, et d'évaluer les indices de sensibilité sur les sorties d'intérêt moyennées. En effet, il n'existe pas, à notre connaissance, de méthodes finalisées d'analyse de sensibilité de modèles stochastiques. Une étude récente propose une méthodologie (basée sur des processus gaussien) pour pallier le problème [\[Marrel et al., 2012\]](#). Il serait peut être nécessaire d'utiliser cette méthode afin d'obtenir des résultats plus fiables.

Discussion-Perspectives

Rappel des objectifs de l'étude

L'objet de cette thèse est le développement de modèles en épidémiologie végétale permettant de prendre en compte divers facteurs pouvant avoir une influence sur la progression de la maladie à l'échelle de la plante mais aussi à plus grande échelle spatiale. Il s'agit aussi de proposer des outils d'inférence statistique bayésienne, dans le cas de modèles complexes et en présence de données incomplètes. L'objectif final est d'étudier la dynamique de la cercosporiose noire du bananier à différentes échelles spatiales et de mieux comprendre sa propagation afin de proposer des moyens de lutte efficace et propre contre cette maladie.

Le chapitre 1 présente le cadre de modélisation et le cadre biologique de la thèse ainsi que les outils utilisés dans les chapitres suivants.

Au chapitre 2 nous avons développé un modèle mécaniste de simulation de la cercosporiose noire du bananier à l'échelle de la plante. Ce modèle utilise la structure du bananier et son évolution dans le temps pour mettre en relation divers paramètres associés à la croissance de la plante sur plusieurs cycles de culture et à la virulence du pathogène au cours de son cycle infectieux. La calibration de ce modèle a nécessité un travail bibliographique et l'utilisation de jeux de données disponibles variés : des données d'expérimentations en laboratoire et des données d'expérimentations au champ. Des expérimentations ont aussi été menées d'une part en Guadeloupe afin d'acquérir de l'information concernant l'architecture de la plante mais aussi en République Dominicaine afin d'acquérir un jeu de données dynamique, à petite échelle, décrivant la surface malade observée sur plusieurs feuilles de plusieurs plantes pendant trois mois. Ce dernier jeu de données a été utilisé pour inférer, au moyen d'une méthode bayésienne, les trois paramètres du modèle les plus influents. La validation du modèle a été effectuée grâce à un jeu de données disponible, à plus grande échelle, d'évaluation du pourcentage de nécroses sur deux cycles de cultures de la plante.

Le chapitre 3 est consacré à l'analyse de sensibilité du modèle SiBaToka développé au chapitre 2. Il s'agit donc d'étudier l'effet des processus associés au développement du pathogène *Mycos-*

phaerella fijiensis. Pour détecter les paramètres les plus influents, deux méthodes sont utilisées : celle de Morris pour laquelle l'espace des paramètres est discrétisé au moyen d'une grille régulière et faisant varier un paramètre à la fois ; la méthode e-FAST qui calcule des indices de sensibilité non linéaires. La première méthode a permis de sélectionner les paramètres les plus influents à étudier plus précisément grâce à la seconde méthode. Les analyses ont été menées sur chacun des modules, en considérant des valeurs de paramètres d'entrée fixées aux valeurs moyennes d'une variété plantain, et sur le modèle complet. Les sorties d'intérêt analysées ont été choisies suivant l'objectif d'évaluation de chacun des modules.

Au chapitre 4, un modèle de propagation spatio-temporelle de la cercosporiose noire à l'échelle de la Martinique est présenté. Ce modèle est markovien et en temps continu. Il postule que le futur état sanitaire d'un patch sain dépend des patches voisins infectés par le biais d'une fonction de dispersion intégrant la distance entre patches. Des covariables environnementales sont également prises en compte dans la probabilité de la future infection. Sur la base des données d'enquêtes effectuées par la FREDON Martinique et de données décrivant le paysage, nous avons pu ajuster ces observations au modèle dans un cadre bayésien en développant un algorithme d'augmentation des données. Cet algorithme construit des dates de première infection avec les contraintes de fenêtre de détection, et peut aussi, sur les patches (ou fenêtres) non observés, effectuer une prévision d'infection et donc proposer une reconstitution de l'histoire de l'invasion de la cercosporiose noire en Martinique.

Dans cette dernière partie, nous discutons l'apport des travaux exposés dans les chapitres 2, 3 et 4 pour la compréhension des dynamiques de la cercosporiose noire et leur gestion, puis nous présentons nos perspectives dans la continuité de ce travail.

Rappel des principaux résultats

Les modèles développés

Modélisation à l'échelle de la plante

La construction du modèle a été menée de façon à ce que tous les paramètres d'entrée dépendent de la variété : les paramètres du cycle épidémiques du pathogène et les paramètres génétiques et agronomiques de la plante. L'outil de simulation proposé permet donc de tester différents scénarios en évaluant l'état sanitaire de la plante en fonction des composantes de résistance choisies et des caractéristiques génétiques et agronomiques définies.

Ce modèle a été conçu de manière à ce que le cycle épidémique et celui de la plante soient décrits le plus précisément possible compte tenu de nos connaissances du pathosystème. Cette précision est en effet nécessaire à une bonne représentation de la réalité.

Nous avons aussi souhaité que la dispersion du pathogène soit modélisée par des fonctions de dispersion, afin de quantifier l'inoculum secondaire par le biais de distances de dispersion.

Pour finir, le modèle a été implémenté de manière à ce que les temps de calculs soient les plus courts possibles et afin de pouvoir obtenir des sorties multiples, notamment les critères d'évaluation de la maladie au champ.

Modélisation à l'échelle d'un territoire

La construction du modèle a été menée de façon à ce que la structure du paysage bananier sur tout le territoire soit prise en compte. Sachant l'effet des conditions climatiques sur la dynamique du pathogène, le modèle prend en compte aussi une variable pluviométrique.

Nous avons souhaité que la dispersion soit modélisée par une fonction dépendante de la distance entre patchs, la taille de ces patchs étant suffisamment petite pour obtenir une évaluation précise de la dispersion.

Les résultats des simulations et des analyses de sensibilité

Le modèle à l'échelle de la plante

L'évaluation de l'état sanitaire d'une feuille ou d'une plante au cours du cycle de culture permet par exemple d'évaluer la proportion de surface malade dans les différents états épidémiques et d'améliorer les connaissances sur l'état sanitaire global au cours des cycles de culture.

Le modèle permet aussi l'évaluation du ratio allo-inoculum/auto-inoculum au cours du cycle, évaluation très difficile à mener expérimentalement (surtout au niveau de la quantification de l'auto-inoculum) mais très informative pour proposer des méthodes de lutte visant à réduire l'auto-inoculum de la plante.

Les analyses de sensibilité ont montré un effet de l'inoculum primaire plus important sur le premier cycle de culture que sur les suivants. Elles ont aussi mis en évidence l'effet majeur de trois composantes de résistance : la vitesse d'extension des lésions, la durée d'incubation et l'efficacité d'infection sur les jeunes feuilles. Nous avons aussi montré l'importance de deux caractères agronomiques : le rythme d'émission foliaire et le nombre de feuilles présentes sur la plante. Ceci permet de proposer ou de prouver la validité des pratiques d'effeuillage et de pratiques agronomiques (telles que la fertilisation et/ou l'irrigation en zones sèches) assurant une émission foliaire rapide pour améliorer l'état sanitaire de la plante. Enfin, les analyses de sensibilité ont montré l'effet de deux des caractères génétiques définis, le nombre de feuilles émises par la plante et la durée entre la floraison et la récolte, ceci pouvant aider à la sélection variétale.

Le modèle à l'échelle régionale

L'évaluation des temps de contamination des patchs a permis de montrer la dépendance de la dynamique invasive à la composition paysagère et climatique du territoire.

Les analyses de sensibilité ont permis de caractériser l'importance des effets des conditions pluviométriques sur le nombre de patchs infectés en début d'épidémie, et de mettre en évidence l'effet de la décroissance de la dispersion en fonction de la distance entre les patchs sur le nombre de patchs infectés tout au long de l'épidémie. Pour finir, les analyses ont aussi montré l'effet du type de paysage sur le nombre de patchs infectés, au moyen d'interactions.

Les données et les résultats d'inférences

Le modèle à l'échelle de la plante

Nous avons pu paramétrer le modèle à partir de données disponibles dans la littérature et par des expérimentations passées. Nous avons cependant choisi de mettre en place une expérimentation en République Dominicaine pour obtenir des informations dynamiques et précises à l'échelle de compartiments de feuilles définis dans le modèle afin de le calibrer, par une méthode bayésienne, et d'obtenir des distributions des trois paramètres les plus influents. Ces données acquises à partir de lésions dessinées deux fois par semaine sur des transparents comprenaient la surface malade et le nombre de lésions. Les nombres de lésions n'ont pas été pris en compte dans l'inférence car la précision de la réalisation de l'expérimentation n'a pas permis de les considérer comme fiables. Pour ces mêmes raisons, l'inférence de la durée d'incubation est certainement surévaluée en raison des observations tardives des petites lésions sur les feuilles.

La méthode d'inférence choisie a été validée et est considérée comme performante pour l'estimation des trois paramètres les plus influents comme le prouvent les validations effectuées sur des jeux de données simulées et la convergence rapide des chaînes. Cependant, des problèmes de convergence apparaissent lorsque l'on augmente le nombre de paramètres. Sachant les informations a priori disponibles sur ces trois paramètres, nous pouvons considérer que les densités a posteriori sont très informatives. Cette méthode permet aussi d'estimer la variance des erreurs du modèle, variance qui tient compte à la fois des erreurs liées aux expérimentations et à celles générées par les simulations du modèle.

Le modèle à l'échelle régionale

Nous disposons de données géoréférencées du parcellaire bananier déclaré sur les trois années de suivi de la maladie, de la couverture du paysage et des hauteurs de pluies dans chacun des 1236 patchs constituant la Martinique. Ces données ont permis de proposer des simulations réalistes de l'invasion de la cercosporiose noire. Cependant compte tenu de la propagation aérienne des spores responsables de la dispersion à longue distance du pathogène, les données de vitesses

et directions du vent auraient certainement amélioré les sorties du modèle.

Le jeu de données constitué du suivi temporel de l'invasion de la maladie à grande échelle sur 83 semaines représente un jeu de données de qualité, mais l'échelle spatiale et temporelle du suivi a nécessairement censuré les données.

Nous avons donc utilisé, dans un cadre bayésien, une méthode d'inférence sur données manquantes, permettant de considérer les temps de colonisation inconnus comme des paramètres supplémentaires. Une démarche "pas-à-pas" a été utilisée afin de valider méthodiquement l'algorithme implémenté.

Nous n'avons aucune information a priori concernant les valeurs des paramètres, et les densités a posteriori semblent très informatives (bien que les résultats finaux ne soient pas encore évalués). Les résultats préliminaires suggèrent que, parmi les 5 types de paysages évalués, les jardins privés sont ceux ayant la plus contribué à l'invasion de la cercosporiose noire en Martinique. Les résultats finaux permettront donc d'évaluer plus précisément l'effet des types de paysage et l'effet des hauteurs de pluie sur la dynamique d'invasion et de définir une fonction de dispersion des ascospores paramétrée à grande échelle spatiale. Les résultats finaux permettront aussi de prédire les dates d'invasion des patchs non renseignés et donc de reconstruire l'histoire de l'invasion de la cercosporiose noire en Martinique.

Perspectives

Résistances

Les perspectives à court terme sont d'améliorer les modèles développés afin de mieux prendre en compte les composantes de résistance de l'hôte au pathogène et de permettre une meilleure approximation de la réalité.

A l'échelle de la plante, les interactions entre les valeurs des paramètres et l'état sanitaire de la plante, et l'interaction entre l'état sanitaire et la dynamique du cycle infectieux semblent intéressantes à évaluer en utilisant des méthodes d'analyse de sensibilité.

L'évolution temporelle des paramètres pourrait être directement intégrée au modèle à différentes échelles temporelles, par exemple en prenant en compte le fait que la quantité de conidies produites pendant la période infectieuse n'est pas constante [Jones, 2000], mais aussi à une échelle temporelle plus large qui permettrait de prendre en compte l'évolution du pathogène et de mieux comprendre les contournements de résistance observés [Fullerton and Olsen, 1995]. En effet, la modélisation de l'évolution des populations pathogènes semble être un sujet de plus en plus développé (Papaïx et al. [2011], Sapoukhina et al. [2009]).

A l'échelle d'un territoire, il semble aussi important de caractériser la résistance dans le modèle en insérant dans un premier temps des durées de latence sexuée.

Dispersion

La dispersion des spores dans les deux modèles a été modélisée par des fonctions de dispersion. Il serait intéressant de comparer plusieurs fonctions afin d'ajuster pour chaque échelle spatiale et chaque type de spores une fonction de dispersion spécifique adaptée.

On pourrait aussi relier la dispersion à une caractéristique climatique (par exemple le vent dans le modèle à l'échelle d'un territoire et l'humidité à l'échelle de la plante) ou bien utiliser des fonctions de dispersion non isotropes (Rieux et al. [2014], Audsley et al. [2005]). Par exemple, à l'échelle de la plante plusieurs auteurs ont modélisé la dispersion des spores entre différentes parties d'une même plante (auto-inoculum) (Audsley et al. [2005], Calonnec et al. [2008], Casadebaig et al. [2012]). Dans ces modèles, les fonctions de dispersion utilisées présentent des caractéristiques spécifiques à la dispersion du pathogène étudié. Audsley et al. [2005] suppose par exemple, que l'inoculum se disperse uniquement des feuilles basses vers les feuilles hautes en raison de l'inefficacité de l'inoculum tombant sur les feuilles basses rapidement sénescentes.

Ces améliorations peuvent permettre une meilleure simulation du pathosystème et peut-être une meilleure prise en compte de la dynamique épidémique à l'échelle de la plante ou la dynamique invasive à l'échelle du territoire.

Création d'un modèle à une échelle intermédiaire

Afin de proposer des méthodes de lutte efficaces, les agronomes s'intéressent d'avantage à la gestion de l'état sanitaire d'une parcelle qu'à celui d'une plante ou d'un territoire. Une perspective serait le développement d'un modèle épidémiologique à une échelle intermédiaire entre la plante et le territoire : à l'échelle d'une parcelle.

Deux propositions sont présentées :

Ce modèle pourrait être basé sur une spatialisation du modèle plante SiBaToka en utilisant des fonctions de dispersion verticales (dispersion inter-plant) mais également horizontales (dispersion entre plants) et prendrait en compte les dynamiques individuelles de chaque plante permettant par exemple d'évaluer des agencements variétaux.

Une seconde manière de décrire la dynamique du pathogène à l'échelle d'une parcelle serait de considérer un modèle à compartiments. Le modèle contiendrait quatre compartiments S : sain, E : latent, C : lésions sporulantes en conidies et A : lésions sporulantes en ascospores. La valeur de chaque compartiment correspondrait à la surface occupée par le stade correspondant, exprimée en nombre de feuilles. Ce modèle pourrait être défini par un système d'équations différentielles.

La spatialisation consisterait à répéter ce modèle dans chaque case d'une grille et à appliquer des fonctions de dispersion pour les spores sexuées et asexuées entre les cases de la grille.

Bibliographie

- Abadie, C., Chilin-Charles, Y., Huat, J., Salmon, F., Pignolet, J., Carlier, J., Lescot, T., Côte, F., and Jenny, C. 2009. New approaches to select durable resistant varieties of banana against mycosphaerella leaf spot diseases. *Acta Horticulturae*, 828 :171–178. [20](#), [86](#), [108](#)
- Abadie, C., El Hadrami, A., Fouré, E., and Carlier, J. 2003. Efficiency and durability of partial resistance components of bananas against black leaf streak disease. In Jacome, L., Lepoivre, P., Marin, D., Ortiz, R., Romero, R., and JV, E., editors, *2th workshop on Mycosphaerella leaf spot diseases of bananas : present status and outlook*, page 317, Montpellier. [86](#), [109](#)
- Amil, A., Heaney, S., Stanger, C., and Shaw, M. W. 2007. Dynamics of QoI sensitivity in *Mycosphaerella fijiensis* in Costa Rica during 2000 to 2003. *Phytopathology*. [17](#), [118](#), [162](#)
- Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., and Daszak, P. 2004. Emerging infectious diseases of plants : pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, 19(10) :535 – 544. [3](#)
- Arneson, P. 1989. A mechanistic computer simulation model of sigatoka. In *Sigatoka leaf spot diseases of bananas*. [19](#)
- Audsley, E., Milne, A., and Paveley, N. 2005. A foliar disease model for use in wheat disease management decision support systems. *Ann. Appl. Biol*, 147 :161–172. [6](#), [174](#)
- Austerlitz, F., Dick, C. W., C., D., Klein, E. K., Oddou-Muratorio, S., Smouse, P. E., and Sork, V. L. 2004. Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology*, 13 :937–954. [5](#), [8](#), [127](#)
- Aylor, D. 1999. Biophysical scaling and the passive dispersal of fungus spores : relationship to integrated pest management strategies. *Agricultural and Forest Meteorology*, 97 :275–292. [4](#)
- Aylor, D. 2003. Spread of plant disease on a continental scale : role of aerial dispersal of pathogens. *Ecology*, 84(8) :1989–1997. [5](#), [7](#)
- Brown, J. and Hovmöller, M. 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297 :537–541. [5](#), [7](#), [117](#)
- Brown, P. E., Chimard, F., Remorov, A., Rosenthal, J. S., and Wang, X. 2013. Statistical inference and computational efficiency for spatial infectious disease models with plantation data. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*. [120](#), [134](#), [165](#)
- Burt, P. J. A., Rosenberg, L. J., Rutter, J., Ramirez, F., and Gonzales O, H. 1999. Forecasting the airborne spread of mycosphaerella fijiensis a cause of black sigatoka disease on banana : estimations of numbers of perithecia and ascospores. *Annals of Applied Biology*, 135(1) :369–377. [71](#), [91](#)

- Burt, P. J. A., Rutter, J., and Ramirez, F. 1998. Airborne spore loads and mesoscale dispersal of the fungal pathogens causing sigatoka diseases in banana and plantain. *Aerobiologia*, 14 :209–214. [17](#), [118](#)
- Calonnec, A., Cartolaro, P., Naulin, J.-M., Bailey, D., and Langlais, M. 2008. A host-pathogen simulation model : powdery mildew of grapevine. *Plant Pathology*, 57(3) :493–508. [6](#), [7](#), [174](#)
- Campolongo, F., Cariboni, J., and Saltelli, A. 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10) :1509–1518. [12](#), [92](#)
- Carlier, J., De Waele, D., Escalant, J.-V., Vézina, A., and Picq, C. 2002. Global evaluation of musa germplasm for resistance to fusarium wilt, mycosphaerella leaf spot diseases and nematodes : in-depth evaluation. Technical report, INIBAP Technical Guidelines, 6. [88](#)
- Carlier, J., Lebrun, M. H., Zapater, M. F., Dubois, C., and Mourichon, X. 1996. Genetic structure of the global population of banana black leaf streak fungus, mycosphaerella fijiensis. *Molecular Ecology*, 5(4) :499–510. [15](#), [18](#)
- Carlier, J., Mourichon, X., and Jones, D. 2000. *Diseases of Banana, Abaca and Enset*, chapter Black leaf streak : the causal agent, pages 48–56. CABI Publishing [New York]. [78](#), [119](#)
- Casadebaig, P., Quesnel, G., Langlais, M., and Faivre, R. 2012. A generic model to simulate air-borne diseases as a function of crop architecture. *PLoS ONE*, 7(11) :1–11. [6](#), [85](#), [174](#)
- Castelan, F., Abadie, C., Hubert, O., Chilin-Charles, Y., de Lapeyre de Bellaire, L., and Chillet, M. 2013. Relation between the severity of sigatoka disease and banana quality characterized by pomological traits and fruit green life. *Crop Protection*, 50(0) :61–65. [95](#)
- Castelan, F., Saraiva, L., Lange, F., de Lapeyre de Bellaire, L., Cordenunsi, B., and Chillet, M. 2012. Effects of black leaf streak disease and sigatoka disease on fruit quality and maturation process of bananas produced in the subtropical conditions of southern Brazil. *Crop Protection*, 35(0) :127–131. [14](#)
- Catterall, S., Cook, A. R., Marion, G., Butler, A., and Hulme, P. E. 2012. Accounting for uncertainty in colonisation times : a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. *Ecography*, 35(10) :901–911. [134](#), [165](#), [166](#)
- Caubel, J. 2012. *Modélisation dynamique et générique de pathosystèmes fongiques aériens. Application à l'étude des impacts du changement climatique sur la rouille brune du blé et le mildiou de la vigne*. PhD thesis, AgroParisTech. [109](#)
- Caubel, J., Launay, M., Lannou, C., and Brisson, N. 2012. Generic response functions to simulate climate-based processes in models for the development of airborne fungal crop pathogens. *Ecological Modelling*, 242(0) :92–104. [6](#)

- Chillet, M., Castelan, F., Abadie, C., Hubert, O., and De Lapeyre De Bellaire, L. 2013. Necrotic leaf removal, a key component of integrated management of *Mycosphaerella* leaf spot diseases to improve the quality of banana : the case of sigatoka disease. *Fruits*, 68 :271–277. [95](#), [110](#)
- Churchill, A. 2011. *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana : progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molecular Plant Pathology*, 12(4) :307–328. [14](#), [15](#), [16](#), [119](#)
- Ciric, C., Ciffroy, P., and Charles, S. 2012. Use of sensitivity analysis to identify influential and non-influential parameters within an aquatic ecosystem model. *Ecological Modelling*, 246 :119–130. [92](#)
- Clark, J. 2005. Why environmental scientist are becoming Bayesians. *Ecology Letters*, 8 :2–14. [9](#)
- Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., and Acutis, M. 2010a. Comparison of sensitivity analysis techniques : A case study with the rice model WARM. *Ecological Modelling*, 221(16) :1897–1906. [92](#)
- Confalonieri, R., Bellocchi, G., Tarantola, S., Acutis, M., Donatelli, M., and Genovese, G. 2010b. Sensitivity analysis of the rice model WARM in Europe : Exploring the effects of different locations, climates and methods of analysis on model sensitivity to crop parameters. *Environmental Modelling & Software*, 25(4) :479–488. [13](#), [85](#), [92](#)
- Cook, A., Marion, G., Butler, A., and Gibson, G. 2007. Bayesian inference for the spatio-temporal invasion of alien species. *Bulletin of Mathematical Biology*, 69 :2005–2025. [8](#), [10](#), [120](#), [127](#), [134](#), [164](#), [165](#)
- Costes, E., Lauri, P. E., Simon, S., and Andrieu, B. 2013. Plant architecture, its diversity and manipulation in agronomic conditions, in relation with pest and pathogen attacks. *Eur J Plant Pathol.* [6](#)
- Craenen, K. and Ortiz, R. 1998. Influence of black sigatoka disease on the growth and yield of diploid and tetraploid hybrid plantains. *Crop Protection*, 17(1) :13–18. [111](#)
- Cressie, N. and Wikle, C. 2011. *Statistics for Spatio-Temporal Data*. Wiley. [133](#)
- Cunniffe, N. J. and Gilligan, C. A. 2010. Invasion, persistence and control in epidemic models for plant pathogens : the effect of host demography. *Journal of The Royal Society Interface*, 7(44) :439–451. [112](#)
- Daley, D. and Vere-Jones, D. 2003. *An Introduction to the Theory of Point Processes*. Springer, 2nd Ed., New York. [7](#), [8](#), [120](#), [164](#)

- De Lapeyre de Bellaire, L., Fouré, E., Abadie, C., and Carlier, J. 2010. Black leaf streak disease is challenging the banana industry. *Fruits*, 65(06) :327–342. [18](#)
- De Wolf, E. D. and Isard, S. A. 2007. Disease cycle approach to plant disease prediction. *Annual Review of Phytopathology*, 45 :203–220. [6](#), [85](#)
- DeJonge, C. K., Ascough II, C. J., Ahmadi, M., Andales, A. A., and Arabi, M. 2012. Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments. *Ecological Modelling*, 231 :113–125. [85](#), [92](#)
- Diggle, P. 2006. Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research*, 15 :325–336. [8](#)
- Drouet, J.-L., Capian, N., Fiorelli, J.-L., Blanfort, V., Capitaine, M., Duretz, S., Gabrielle, B., Martin, R., Lardy, R., Cellier, P., and Soussana, J.-F. 2011. Sensitivity analysis for models of greenhouse gas emissions at farm level. Case study of n_2o emissions simulated by the CERES-EGC model. *Environmental Pollution*, 159 :3156–3161. [92](#)
- Eikemo, H., Gadoury, D., Villalta, O., Creemers, P., and Seem, R. C. and Stensvand, A. 2011. Evaluation of six models to estimate ascospore maturation in *Venturia pyrina*. *Plant Disease*, 95(3) :279–284. [80](#)
- Fabre, F., Bruchou, C., Palloix, A., and Moury, B. 2009. Key determinants of resistance durability to plant viruses : Insights from a model linking within- and between-host dynamics. *Virus Research*, 141(2) :140–149. Plant Virus Epidemiology : Controlling epidemics of emerging and established plant viruses - the way forward. [111](#)
- Facon, B. and David, P. 2006. Metapopulation dynamics and biological invasions : a spatially explicit model applied to a freshwater snail. *The American naturalist*, 168(6) :769–783. [7](#)
- Faivre, R., Ioos, B., Mahévas, S., Makowski, D., and Monod, H. 2013. *Analyse de sensibilité et exploration de modèles. Application aux sciences de la nature et de l'environnement*. Quae Editions. [13](#), [85](#), [92](#)
- Fewster, R. M. 2003. A spatiotemporal stochastic process model for species spread. *Biometrics*, 59(3) :640–649. [8](#), [119](#)
- Filipe, J. and Maule, M. 2004. Effects of dispersal mechanisms on spatio-temporal development of epidemics. *Journal of Theoretical Biology*, 226 :125–141. [119](#)
- Filipe, J., Otten, W., Gibson, G., and Gilligan, C. 2004. Inferring the dynamics of a spatial epidemic from time-series data. *Bulletin of Mathematical Biology*, 66 :373–391. [120](#)

- Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. *Nature*, 484(7393) :186–194. [117](#)
- Fortin, M.-J., Boots, B., Csillag, F., and Rimmel, T. 2003. On the role of spatial stochastic models in understanding landscape indices in ecology. *Oikos*, 102(1) :203–212. [164](#)
- Fortune, M. P., Gosine, S., Chow, S., Dilbar, A., St Hill, A., Gibbs, H., and Rambaran, N. 2005. First report of black sigatoka disease (causal agent *Mycosphaerella fijiensis*) from trinidad. *Plant Pathology*, 54(2) :246–246. [18](#), [117](#)
- Fouré, E. 1982. Les cercosporioses du bananier et leurs traitements. comportement des variétés. etude de la sensibilité variétale des bananiers et plantain à mycosphaerella fijiensis morelet au Gabon (maladie des raies noires). *Fruits*, 37(12) :749–766. [71](#), [80](#), [91](#)
- Fouré, E., Grisoni, M., and Zurfluh, R. 1984. Les cercosporioses du bananier et leurs traitements. comportement des variétés. *Fruits*, 39(6) :365–378. [71](#), [91](#), [108](#), [111](#)
- Fouré, E., Mourichon, X., and Jones, D. 2000. *Diseases of Banana, abaca and Enset. Evaluating germplasm for reaction to black leaf streak.*, chapter 2, pages 62–72. CABI Publishing. [17](#)
- Frezal, L., Robert, C., Bancal, M.-O., and Lannou, C. 2009. Local dispersal of *Puccinia triticina* and wheat canopy structure. *Phytopathology*, 99 :1216–1224. [109](#)
- Fullerton, R. A. and Olsen, T. 1995. Pathogenic variability in *Mycosphaerella fijiensis* morelet, cause of black sigatoka in banana and plantain. *New Zealand Journal of Crop and Horticultural Science*, 23 :39–48. [173](#)
- Gauhl, F. 1994. *Epidemiology and Ecology of Black Sigatoka (Mycosphaerella fijiensis Morelet) on Plantain and Banana (Musa spp.) in Costa Rica, Central America*. inibap. [17](#), [71](#), [91](#)
- Gelman, A., Carlin, J., Stern, H., and Rubin, D. 2004. *Bayesian data analysis*. Chapman and Hall-CRC, 2nd edition. [10](#), [137](#)
- Gelman, A. and Rubin, D. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7 :457–511. [11](#), [134](#)
- Gibson, G. 1997. Markov chain monte carlo methods for fitting spatiotemporal stochastic models in plant epidemiology. *Royal statistical society, Appl. Statist*, 42 :215–233. [120](#)
- Gibson, G., Streftaris, G., and Zachary, S. 2010. Generalised data augmentation and posterior inferences. *Journal of statistical planning and inference*, 141 :156–171. [120](#)

- Gilligan, C. A. 2008. Sustainable agriculture and plant diseases : an epidemiological perspective. *Philosophical Transactions of the Royal Society B - Biological Sciences*, 363(1492) :741–759. [85](#), [119](#)
- Hartig, E., Calabrese, J., Reineking, B., Wiegand, T., and Huth, A. 2011. Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, 14 :816–827. [120](#)
- Hastings, A., Cuddington, K., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B. A., Moore, K., Taylor, C., and Thomson, D. 2005. The spatial spread of invasions : new developments in theory and evidence. *Ecology Letters*, 8(1) :91–101. [128](#)
- Henderson, J., Pattemore, J., Porchun, S., Hayden, H., Van Brunschot, S., Grice, K., Peterson, R., Thomas-Hall, S., and Aitken, E. 2006. Black Sigatoka disease : new technologies to strengthen eradication strategies in Australia. *Australian Plant Pathology*, 35 :181–193. [120](#), [163](#)
- Höhle, M., Jorgensen, E., and O'Neill, P. D. 2005. Inference in disease transmission experiments by using stochastic epidemic models. *Journal of the Royal Statistical Society Series C*, 54(2) :349–366. [164](#)
- Hong, J. and Hwang, B. 1998. Influence of inoculum density, wetness duration, plant age, inoculation method, and cultivar resistance on infection of pepper plants by *Colletotrichum coccodes*. *Plant Disease*, 82 :1079–1083. [109](#)
- Ioos, R., Hubert, J., Abadie, C., Duféal, D., Opdebeeck, G., and Iotti, J. 2011. First report of black sigatoka disease in banana caused by *Mycosphaerella fijiensis* on Martinique island. *Plant Disease*, 95(3) :359–359. [18](#), [117](#)
- Iooss, B. 2010. Review of global sensitivity analysis of numerical models. *Journal de la Société Française de Statistique*. [12](#), [13](#), [92](#), [111](#)
- Irish, B., Goenaga, R., Rios, C., Chavarria-Carvajal, J., and Ploetz, R. 2013. Evaluation of banana hybrids for tolerance to black leaf streak (*Mycosphaerella fijiensis* morelet) in Puerto Rico. *Crop Protection*, 54 :229–238. [70](#), [89](#), [90](#)
- Jeger, M. 2004. Analysis of disease progress as a basis for evaluating disease management practice. *Ann. Rev. Phytopathol.*, 42 :61–82. [6](#), [85](#)
- Jones, D. 2000. *Diseases of banana, abaca and enset*. CABI publishing. [4](#), [14](#), [18](#), [71](#), [80](#), [85](#), [91](#), [117](#), [118](#), [173](#)
- Lamboni, M., Makowski, D., Lehuger, S., Gabrielle, B., and Monod, H. 2009. Multivariate global sensitivity analysis for dynamic crop models. *Field Crops Research*, 113(3) :312–320. [85](#), [111](#)

- Lannou, C., Soubeyrand, S., Frezal, L., and ChadÅ“uf, J. 2008. Autoinfection in wheat leaf rust epidemics. *New Phytologist*, 177(4) :1001–1011. [5](#), [109](#)
- Lassoudière, A. 2006. *Le bananier et sa culture*. Editions Quae. [18](#), [70](#), [71](#), [89](#)
- Lassoudière, A. 2012. *Le bananier, Un siècle d'innovations techniques*. Editions Quae. [70](#), [89](#), [90](#)
- Leclerc, M., DorÃ©, T., Gilligan, C. A., Lucas, P., and Filipe, J. A. N. 2014. Estimating the delay between host infection and disease (incubation period) and assessing its significance to the epidemiology of plant diseases. *PLoS ONE*, 9(1) :1–15. [108](#), [111](#)
- Lô Pelzer, E., Aubertot, J., Bousset, L., Salam, M., and Jeuffroy, M. 2010a. SIPPOM-WOSR : A simulator for integrated pathogen population management of phoma stem canker on winter oilseed rape : li sensitivity analysis. *Field Crops Research*, 118(1) :82–93. [85](#), [111](#)
- Lô Pelzer, E., Bousset, L., Jeuffroy, M., Salam, M., Pinochet, X., Boillot, M., and Aubertot, J. 2010b. SIPPOM-WOSR : A simulator for integrated pathogen population management of phoma stem canker on winter oilseed rape : I. description of the model. *Field Crops Research*, 118(1) :73–81. [5](#)
- Long, J. A., Robertson, C., Nathoo, F. S., and Nelson, T. A. 2012. A bayesian space time model for discrete spread processes on a lattice. *Spatial and Spatio-temporal Epidemiology*, 3(2) :151–162. [8](#), [119](#)
- Lurette, A., Touzeau, S., Lamboni, M., and Monod, H. 2009. Sensitivity analysis to identify key parameters influencing salmonella infection dynamics in a pig batch. *Journal of Theoretical Biology*, 258(1) :43–52. [111](#)
- Makowski, D., Hillier, J., Wallach, D., Andrieu, B., and Jeuffroy, M.-H. 2006a. *Parameter estimation for crop models*, chapter 3, pages 101–149. Elsevier. [8](#), [9](#)
- Makowski, D., Naud, C., Jeuffroy, M.-H., Barbottin, A., and Monod, H. 2006b. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. *Reliability Engineering & System Safety*, 91 :1142–1147. [13](#), [92](#), [142](#)
- Marin, D. H., Romero, R. A., Guzman, M., and Sutton, T. B. 2003. Black sigatoka : An increasing threat to banana cultivation. *Plant Disease*, 87(3) :208–222. [14](#), [17](#), [18](#), [110](#)
- Marrel, A., Ioos, B., Da Veiga, S., and Ribatet, M. 2012. Global sensitivity analysis of stochastic computer models with joint metamodels. *Stat Comput*, 22 :833–847. [166](#)
- Matthews, L. and Haydon, D. 2007. Introduction. cross-scale influences on epidemiological dynamics : from genes to ecosystems. *Journal of the Royal Society Interface*, 4(763-765). [5](#)

- McKinley, T. J., Ross, J. V., Deardon, R., and Cook, A. R. 2014. Simulation-based Bayesian inference for epidemic models. *Computational Statistics & Data Analysis*, 71(0) :434–447. [165](#)
- Meentemeyer, R. K., Cunliffe, N. J., Cook, A. R., Filipe, J. A. N., Hunter, R. D., Rizzo, D. M., and Gilligan, C. A. 2011. Epidemiological modeling of invasion in heterogeneous landscapes : spread of sudden oak death in california (1990 à 2030). *Ecosphere*, 2(2). [8](#), [120](#), [166](#)
- Milne, A., Paveley, N., Audsley, E., and Livermore, P. 2003. A wheat canopy model for use in disease management decision support systems. *Ann. appl. Biol.*, 143 :265–274. [6](#)
- Mobambo, K., Gauhl, F., Pasberg-Gauhl, C., and Zuofa, K. 1996. Season and plant age effect evaluation of plantain for response to black sigatoka disease. *Crop Protection*, 15(7) :609–614. [109](#), [111](#)
- Mobambo, K., Gauhl, F., Vuylsteke, D., Ortiz, R., Pasberg-Gauhl, C., and Swennen, R. 1993. Yield loss in plant from black sigatoka leaf spot and field performance of resistant hybrids. *Field Crops Research*, 35 :35–42. [81](#), [111](#)
- Mobambo, K., Zuofa, K., Gauhl, F., Adeniji, O., and Paseberg-Gauhl, C. 1994. Effect of soil fertility on host response to black leaf streak of plantain (musa spp. AAB group) under traditional farfarm systems in southeastern Nigeria. *International journal of pest management*, 40(1) :75–80. [110](#)
- Monod, H., Naud, C., and Makowski, D. 2006. *Uncertainty and sensitivity analysis for crop models*, chapter 3, page 55. Elsevier. [93](#), [111](#)
- Morris, D. J., Speirs, D. C., Cameron, A. I., and Heath, M. R. 2014. Global sensitivity analysis of an end-to-end marine ecosystem model of the north sea : Factors affecting the biomass of fish and benthos. *Ecological Modelling*, 273 :251–263. [92](#)
- Morris, M. D. 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2) :162–174. [92](#), [142](#)
- Mundt, C. 2009. Importance of autoinfection to the epidemiology of polycyclic foliar disease. *Phytopathology*, 99(10) :1116–1120. [5](#)
- Mundt, C. and Sackett, K. 2012. Spatial scaling relationship for spread of disease caused by a wind-dispersed plant pathogen. *Ecosphere*, 3(3). [128](#)
- Mundt, C., Sackett, K., and Wallace, L. 2011. Landscape heterogeneity and disease spread : experimental approaches with a plant pathogen. *Ecological Applications*, 21(2) :321–328. [163](#)
- Neri, F. M., Bates, A., Fuchtbauer, W. S., Perez-Reche, F. J., Taraskin, S. N., Otten, W., Bailey, D. J., and Gilligan, C. A. 2011. The Effect of Heterogeneity on Invasion in Spatial Epidemics : From Theory to Experimental Evidence in a Model System. *Plos Computational Biology*, 7(9). [119](#)

- Noupadja, P., Tomekpe, K., and Youmbi, E. 2007. Évaluation d'hybrides tétraploïdes de bananiers plantains (*musa spp.*) résistants à la maladie des raies noires créés au cameroun. *Fruits*, 62 :77–88. [110](#)
- O'Hara, R. B., Arjas, E., Toivonen, H., and Hanski, I. 2002. Bayesian analysis of metapopulation data. *Ecology*, 83(9) :2408–2415. [10](#)
- O'Neill, P. 2002. A tutorial introduction to bayesian inference for stochastic epidemics models using Markov chain monte carlo methods. *Mathematical Biosciences*, 180 :103–114. [120](#)
- Ostfeld, R. S., Glass, G. E., and Keesing, F. 2005. Spatial epidemiology : an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution*, 20(6) :328 – 336. [8](#)
- Papaïx, J. 2011. *Structure du paysage agricole et risque épidémique, une approche démo-génétique*. PhD thesis, AgroParisTech. [81](#)
- Papaïx, J., Goyeau, H., Du Cheyron, P., Monod, H., and Lannou, C. 2011. Influence of cultivated landscape composition on variety resistance : an assessment based on wheat leaf rust epidemics. *New Phytologist*. [6](#), [20](#), [173](#)
- Parent, E. and Bernier, J. 2007. *Le raisonnement bayésien. Modélisation et inférence*. Springer. [9](#)
- Parnell, S., Gottwald, T. R., Gilks, W., and Van den Bosch, F. 2012. Estimating the incidence of an epidemic when it is first discovered and the design of early detection monitoring. *Journal of Theoretical Biology*, 305 :30–36. [163](#), [164](#)
- Parnell, S., Gottwald, T. R., Gilligan, C. A., Cunniffe, N. J., and van den Bosch, F. 2010. The effect of landscape pattern on the optimal eradication zone of an invading epidemic. *Phytopathology*, 100(7) :638–644. [119](#), [120](#)
- Parnell, S., Gottwald, T. R., Riley, T., and van den Bosch, F. 2013. A generic risk-based surveying method for invading plant pathogens. *Ecological Applications*, 24(4) :779–790. [163](#)
- Plantegenest, M., Le May, C., and Fabre, F. 2007. Landscape epidemiology of plant diseases. *Journal of The Royal Society Interface*, 4(16) :963–972. [8](#), [119](#), [164](#)
- R Development Core Team 2008. *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. [11](#), [24](#), [93](#), [143](#)
- Rasmussen, J. G., Møller, J., Aukema, B. H., Raffa, K. F., and Zhu, J. 2007. Continuous time modelling of dynamical spatial lattice data observed at sparsely distributed times. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 69(4) :701–713. [120](#)
- Real, L. A. and Biek, R. 2007. Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes. *Journal of the Royal Society Interface*, 4(16) :935–948. [8](#)

- Richter, G., Acutis, M., Trevisiol, P., Latiri, K., and Confalonieri, R. 2010. Sensitivity analysis for a complex crop model applied to durum wheat in the mediterranean. *European Journal of Agronomy*, 32 :127–136. [85](#), [92](#)
- Rieux, A. 2011. *Etude des processus de dispersion et des flux géniques chez un champignon pytopathogène : le cas de Mycosphaerella fijiensis à l'échelle d'un bassin de production Camerounais*. PhD thesis, Montpellier SupAgro. [16](#)
- Rieux, A., Lenormand, T., Carlier, J., de Lapeyre de Bellaire, L., and Ravigné, V. 2013. Using neutral cline decay to estimate contemporary dispersal : a generic tool and its application to a major crop pathogen. *Ecology Letters*, 16(6) :721–730. [71](#), [91](#)
- Rieux, A., Soubeyrand, S., Bonnot, F., Klein, E. K., Ngando, J. E., Mehl, A., Ravigné, V., Carlier, J., and de Lapeyre de Bellaire, L. 2014. Long-distance wind-dispersal of spores in a fungal plant pathogen : estimation of anisotropic dispersal kernels from an extensive field experiment. *Plos One*. [17](#), [80](#), [118](#), [162](#), [174](#)
- Robert, C. and Casella, G. 2010. *Monte Carlo Statistical Methods*. second ed. Springer. [9](#), [134](#)
- Robert, S., Ravigné, V., Zapater, M.-E., Abadie, C., and Carlier, J. 2012. Contrasting introduction scenarios among continents in the worldwide invasion of the banana fungal pathogen mycosphaerella fijiensis. *Molecular Ecology*, pages 11–12. [18](#), [117](#)
- Robinson, J. 2003. *Bananas and Plantains*. CAB International. [70](#), [89](#), [90](#)
- Ronce, O. 2007. How does it feel to be like a rolling stone ? ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, 38(1) :231–253. [5](#)
- Ross, J. 2011. Invasion of infectious diseases in finite homogeneous populations. *Journal of Theoretical Biology*, 289(0) :83–89. [8](#), [119](#)
- Rossi, V., Caffi, T., Giosuè, S., and Bugiani, R. 2008. A mechanistic model simulating primary infections of downy mildew in grapevine. *Ecological Modelling*, 212 :480–491. [6](#)
- Rutter, J., Burt, P. J. A., and Ramirez, F. 1998. Movement of *Mycosphaerella fijiensis* spores and sigatoka disease development on plant close to an inoculum source. *Aerobiologia*, 14 :201–208. [17](#), [118](#)
- Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., Commission, E., and of Ispra (I), J. R. C. 2006. Sensitivity analysis practices : Strategies for model-based inference. *Reliability Engineering & System Safety*, 91 :1109–1125. [111](#)
- Saltelli, A., Tarantola, S., and Chan, K. 1999. A quantitative model independent method for global sensitivity analysis of model output. *Technometrics*, 41 :39–56. [13](#), [92](#), [142](#)

- Sapoukhina, N., Durel, C.-E., and Le Cam, B. 2009. Spatial deployment of gene-for-gene resistance governs evolution and spread of pathogen populations. *Theoretical Ecology*, 2(4) :229–238. [173](#)
- Schreiber, S. J. and Lloyd-Smith, J. 2009. Invasion dynamics in spatially heterogeneous environments. *American Naturalist*, 174(4) :490–505. [8](#), [120](#)
- Schurr, F. M., Steinitz, O., and Nathan, R. 2008. Plant fecundity and seed dispersal in spatially heterogeneous environments : models, mechanisms and estimation. *Journal of Ecology*, 96(4) :628–641. [8](#)
- Song, X., Bryan, B. A., Almeida, A. C., Paul, K. I., Zhao, G., and Ren, Y. 2013. Time-dependent sensitivity of a process-based ecological model. *Ecological Modelling*, 265(0) :114–123. [86](#)
- Song, X., Bryan, B. A., Paul, K. I., and Zhao, G. 2012. Variance-based sensitivity analysis of a forest growth model. *Ecological Modelling*, 247(0) :135 – 143. [86](#), [92](#)
- Soubeyrand, S., Held, L., Höhle, M., and Sache, I. 2008. Modelling the spread in space and time of an airborne plant disease. *The royal statistical society, Appl. Statist*, 57(3) :253–272. [7](#)
- Soubeyrand, S., Laine, A., Hanski, I., and Penttinen, A. 2009. Spatiotemporal structure of host-pathogen interactions in a metapopulation. *The American Naturalist*, 174(3). [7](#)
- Stover, R. 1980. Sigatoka leaf spots of bananas and plantains. *Plant disease*, 64(8). [119](#)
- Stover, R. and Simmonds, N. 1987. *Bananas*. Longman Scientific & Technical, third edition. [70](#), [89](#), [90](#)
- Strona, G. 2015. A spatially explicit model to investigate how dispersal/colonization tradeoffs between short and long distance movement strategies affect species ranges. *Ecological Modelling*, 297(0) :80 – 85. [163](#)
- Sumner, T., Shephard, E., and Bogle, I. D. L. 2012. A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modelling. *Journal of the Royal Society Interface*, 9 :2156–2166. [142](#)
- Tanner, M. A. and Wong, W. 2010. From EM to Data Augmentation : The emergence of MCMC Bayesian computation in the 1980s. *Statistical Science*, 25(4) :506–516. [120](#)
- Tixier, P., Malezieux, E., and Dorel, M. 2004. Simba-pop : a cohort population model for long-term simulation of banana crop harvest. *Ecological Modelling*, 180 :407–417. [6](#), [19](#)
- Tixier, P., Risède, J.-M., Dorel, M., and Malézieux, E. 2006. Modelling population dynamics of banana plant-parasitic nematodes : A contribution to the design of sustainable cropping systems. *Ecological Modelling*, 198(3â4) :321–331. [6](#)

- Vaillant, J., Puggioni, G., Waller, L., and Daugrois, J. 2010. A spatio-temporal analysis of the spread of sugarcane yellow leaf virus. *Journal of time series analysis*. 7, 120, 164
- van den Berg, F., Gaucel, S., Lannou, C., Gilligan, C. A., and van den Bosch, F. 2013. High levels of auto-infection in plant pathogens favour short latent periods : a theoretical approach. *Evolutionary Ecology*, 27 :409–428. 110
- van Maanen, A. and Xu, X.-M. 2003. Modelling plant disease epidemics. *European Journal of Plant Pathology*, 109(669-682). 80
- Varella, H., Guérif, M., and Buis, S. 2010. Global sensitivity analysis measures the quality of parameter estimation : The case of soil parameters and a crop model. *Environmental Modelling & Software*, 25(3) :310–319. 13, 92, 142
- Véra, C. 2004. *Modèles linéaires mixtes multiphasiques pour l'analyse de données longitudinales - Application à la croissance des plantes*. PhD thesis, Université Montpellier II. 6
- Wang, J., Li, X., Lu, L., and Fang, F. 2013. Parameter sensitivity analysis of crop growth models based on the extended Fourier Amplitude Sensitivity Test method. *Environmental Modelling & Software*, 48 :171–182. 13, 85, 92
- Wilocquet, L. and Savary, S. 2004. An epidemiological simulation model with three scales of spatial hierarchy. *Phytopathology*, 94 :883–891. 7
- Zhao, G., Bryan, B. A., and Song, X. 2014. Sensitivity and uncertainty analysis of the APSIM-wheat model : Interactions between cultivar, environmental, and management parameters. *Ecological Modelling*, 279 :1–11. 92

Annexe A

Analyse de sensibilité du modèle SiBaToka

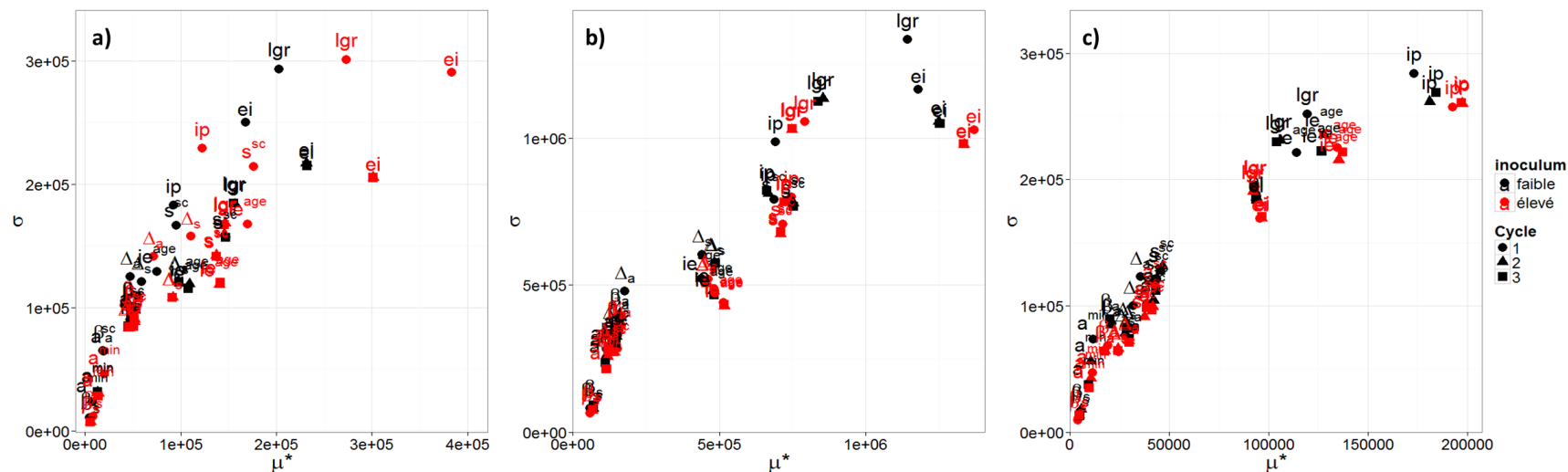


FIGURE A.1 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : nombre de lésions en incubation sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte.

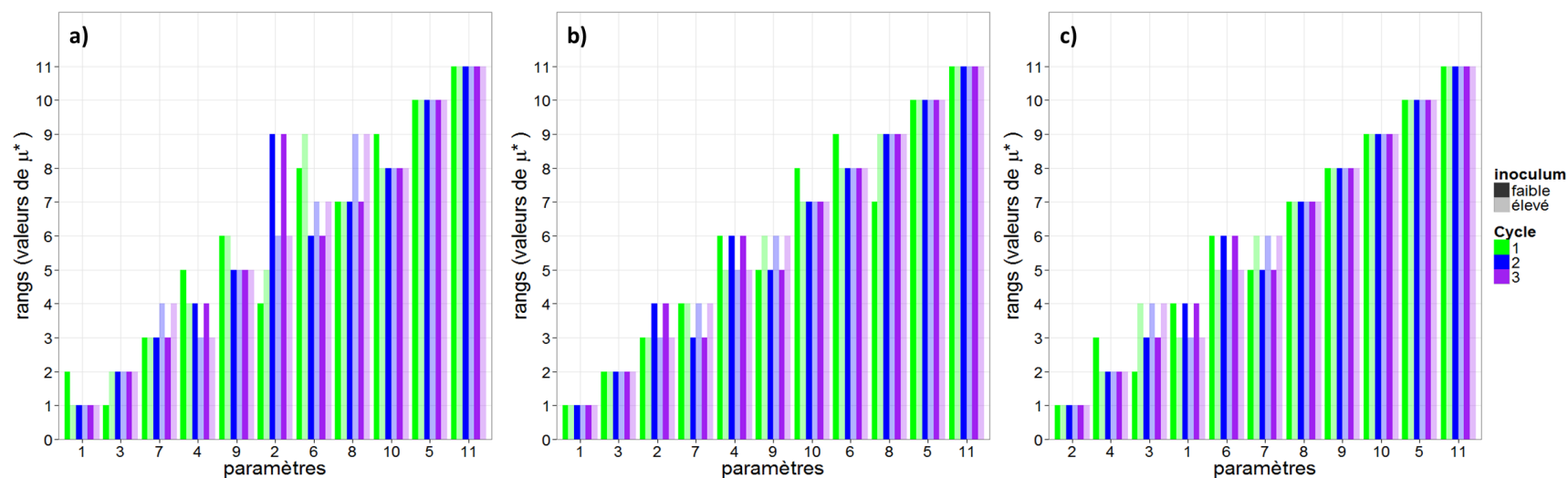


FIGURE A.2 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : nombre de lésions en incubation sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 11 désigne la valeur la plus faible. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s .

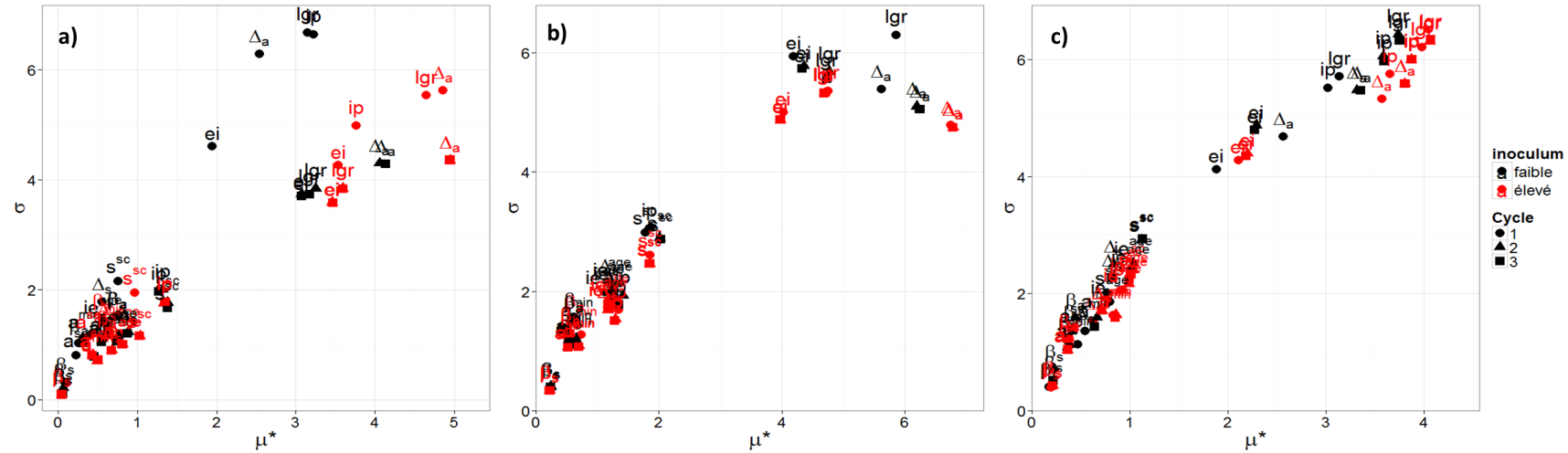


FIGURE A.3 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : pourcentage de surface conidienne de la plante sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte.

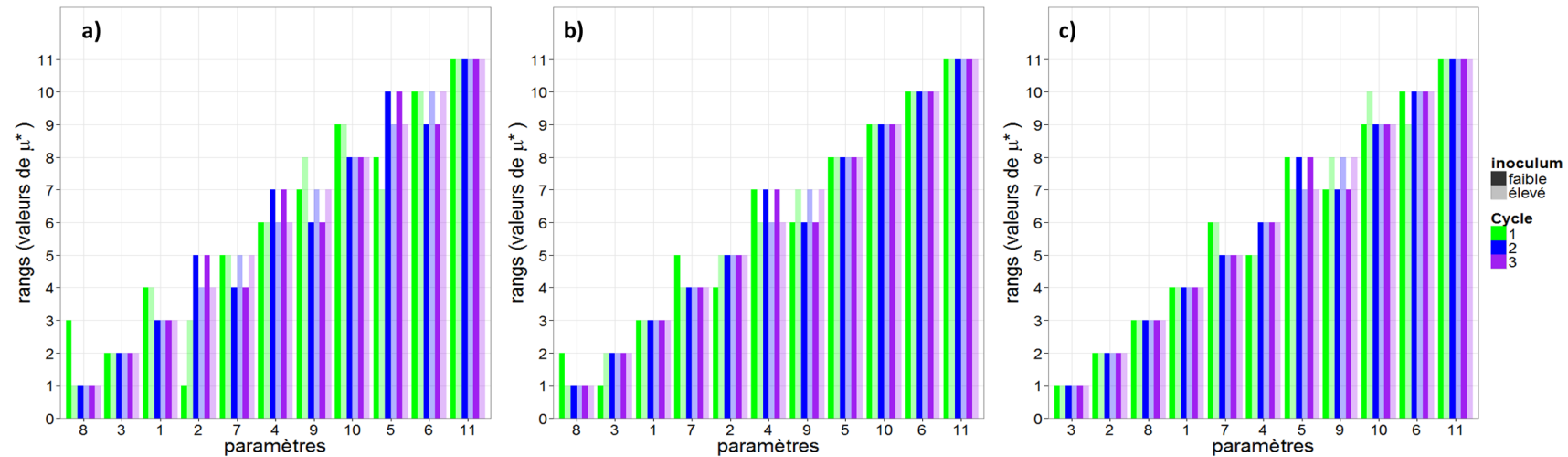


FIGURE A.4 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : pourcentage de surface conidienne de la plante sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 11 désigne la valeur la plus faible. Numérotation des paramètres : 1: ei , 2: ip , 3: lgr , 4: ei^{age} , 5: a^{min} , 6: sc_a , 7: sc_s , 8: Δ_a , 9: Δ_s , 10: β_a , 11: β_s .

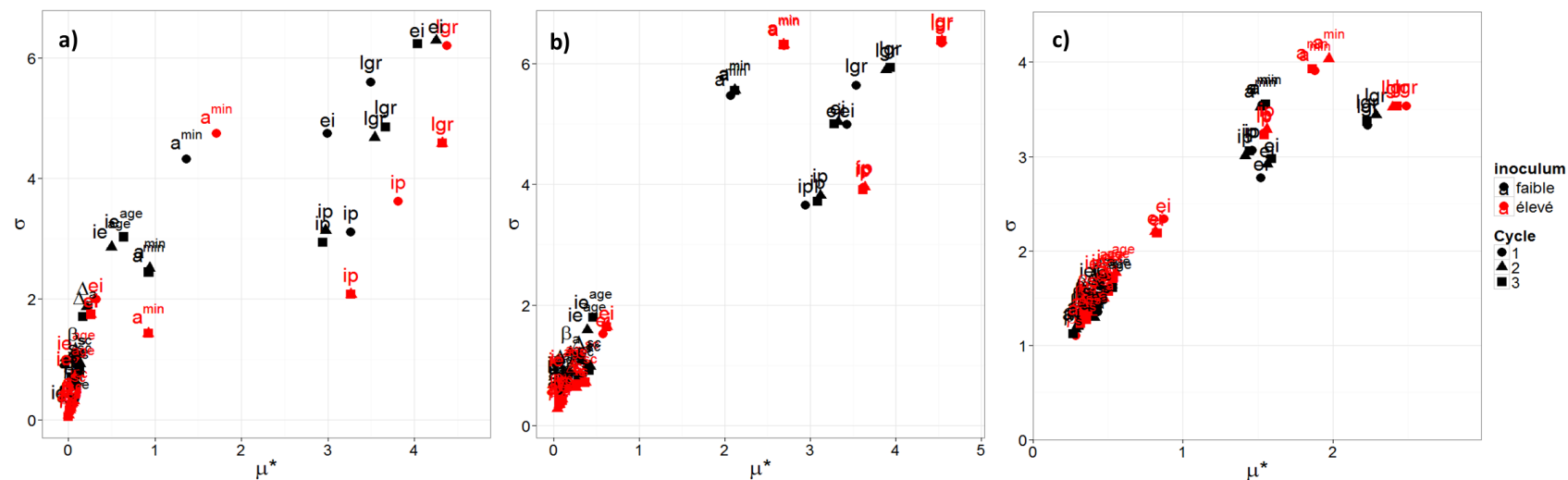


FIGURE A.5 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : PJFM sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte.

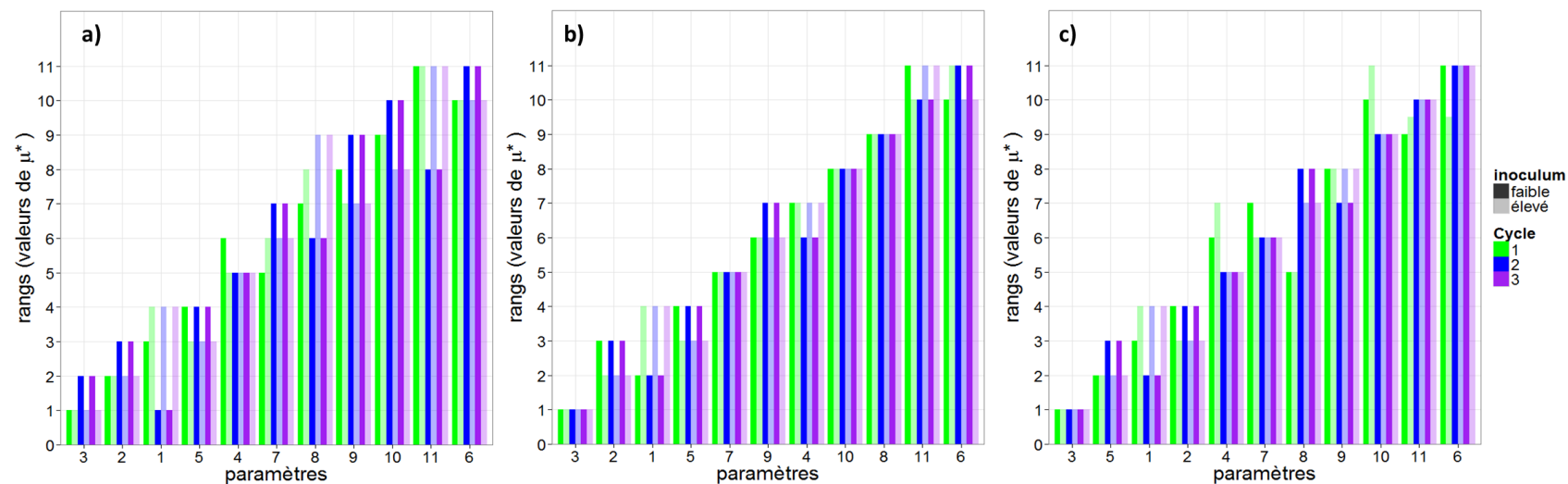


FIGURE A.6 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : PJFM sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 11 désigne la valeur la plus faible. Numérotation des paramètres : 1: ei , 2: ip , 3: lgr , 4: ei^{age} , 5: a^{min} , 6: sc_a , 7: sc_s , 8: Δ_a , 9: Δ_s , 10: β_a , 11: β_s .

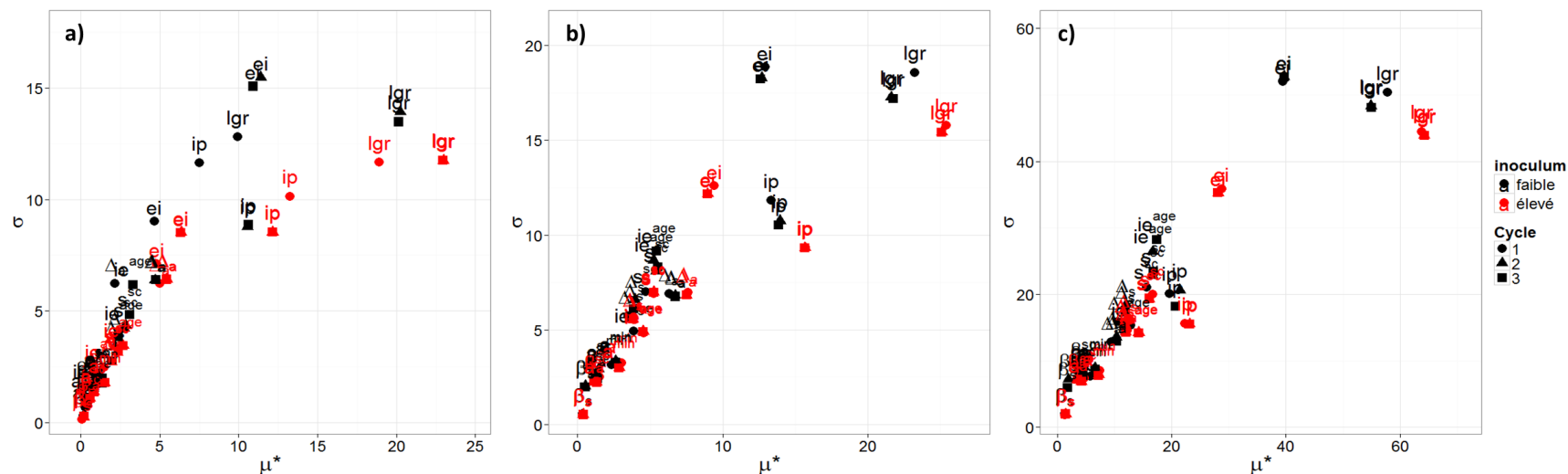


FIGURE A.7 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : pourcentage de nécroses de la plante sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte.

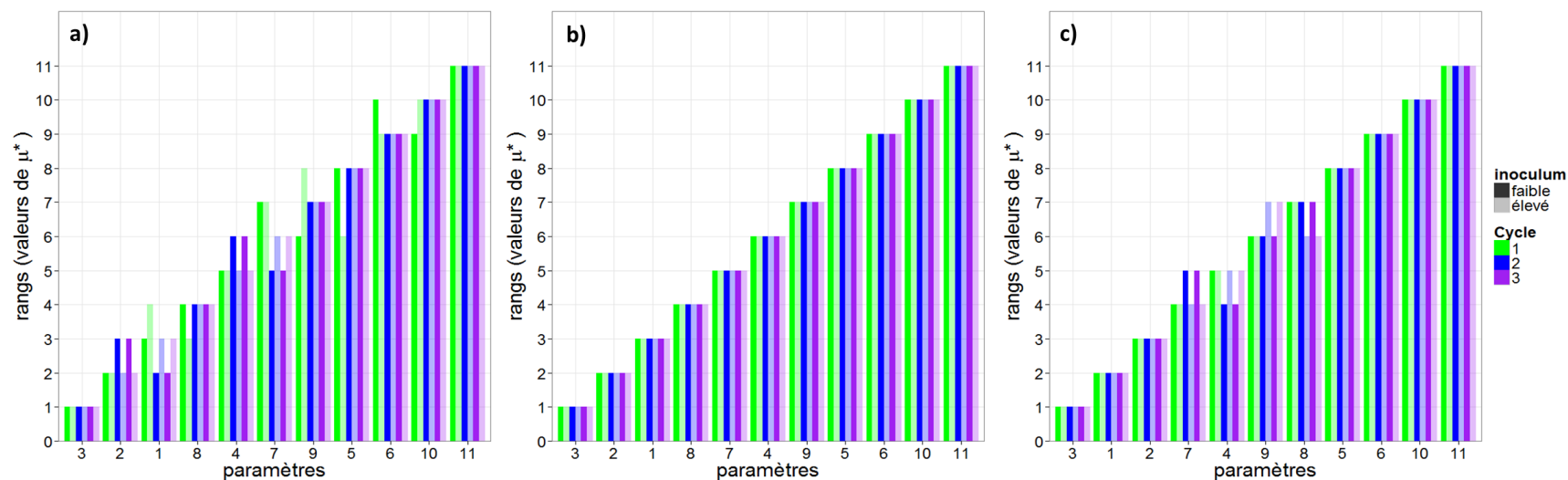


FIGURE A.8 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : pourcentage de nécroses de la plante sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 11 désigne la valeur la plus faible. Numérotation des paramètres : 1: ei , 2: ip , 3: lgr , 4: ei^{age} , 5: a^{min} , 6: sc_a , 7: sc_s , 8: Δ_a , 9: Δ_s , 10: β_a , 11: β_s .

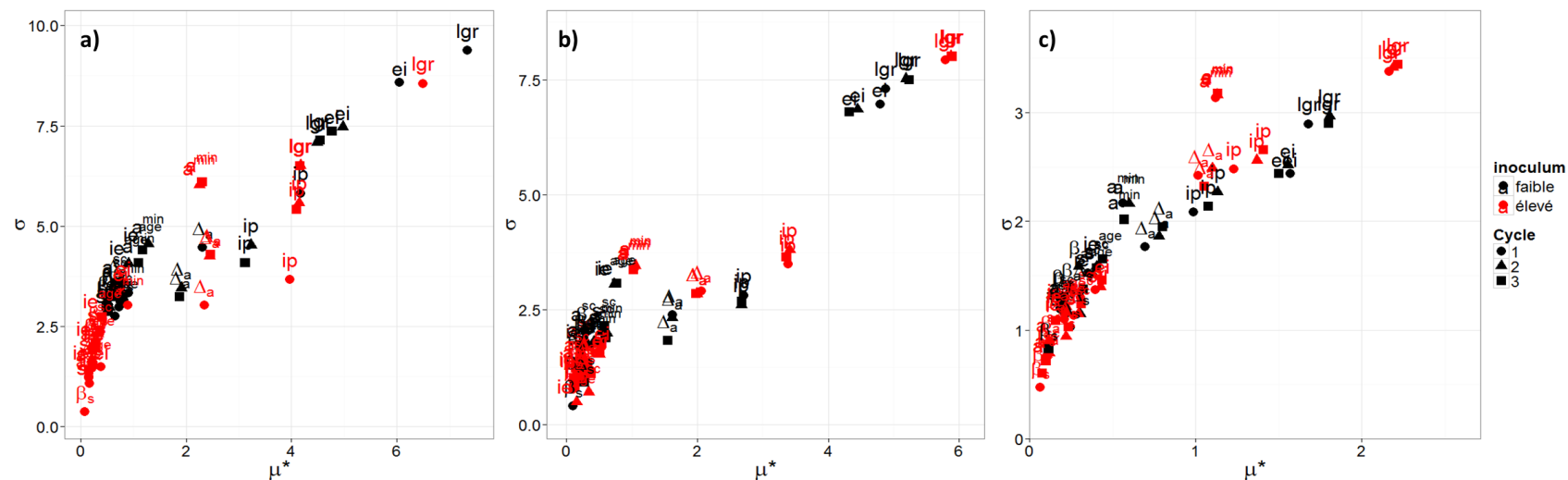


FIGURE A.9 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : PJFN sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte.

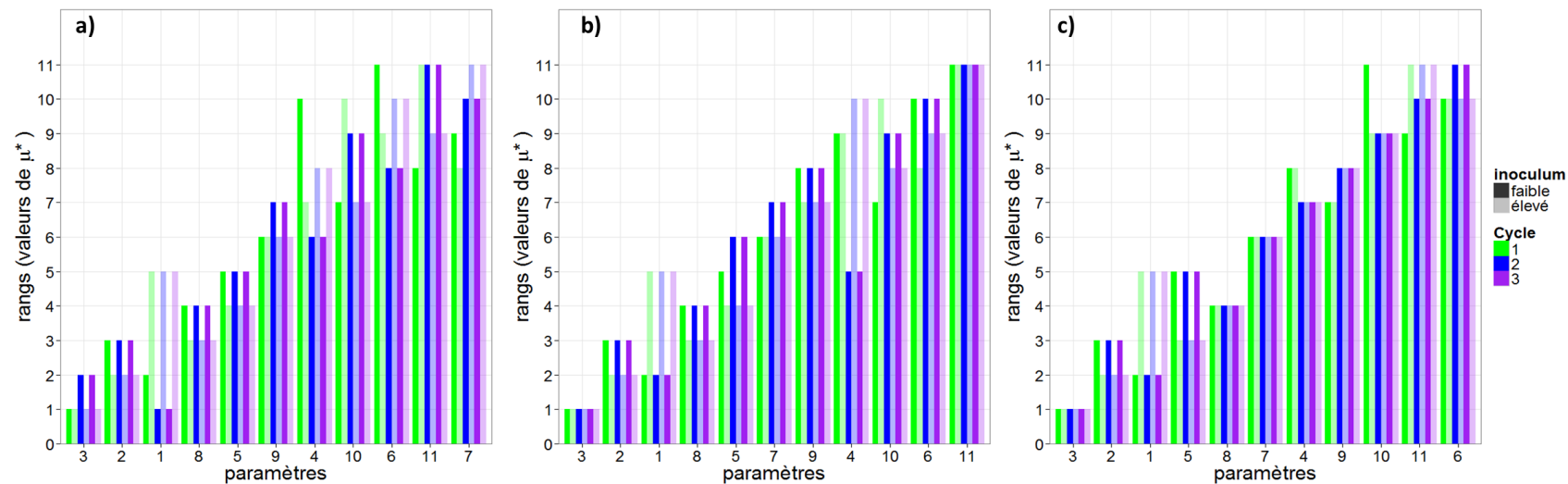


FIGURE A.10 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : PJFN sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 11 désigne la valeur la plus faible. Numérotation des paramètres : 1: ei , 2: ip , 3: lgr , 4: ei^{age} , 5: a^{min} , 6: sc_a , 7: sc_s , 8: Δ_a , 9: Δ_s , 10: β_a , 11: β_s .

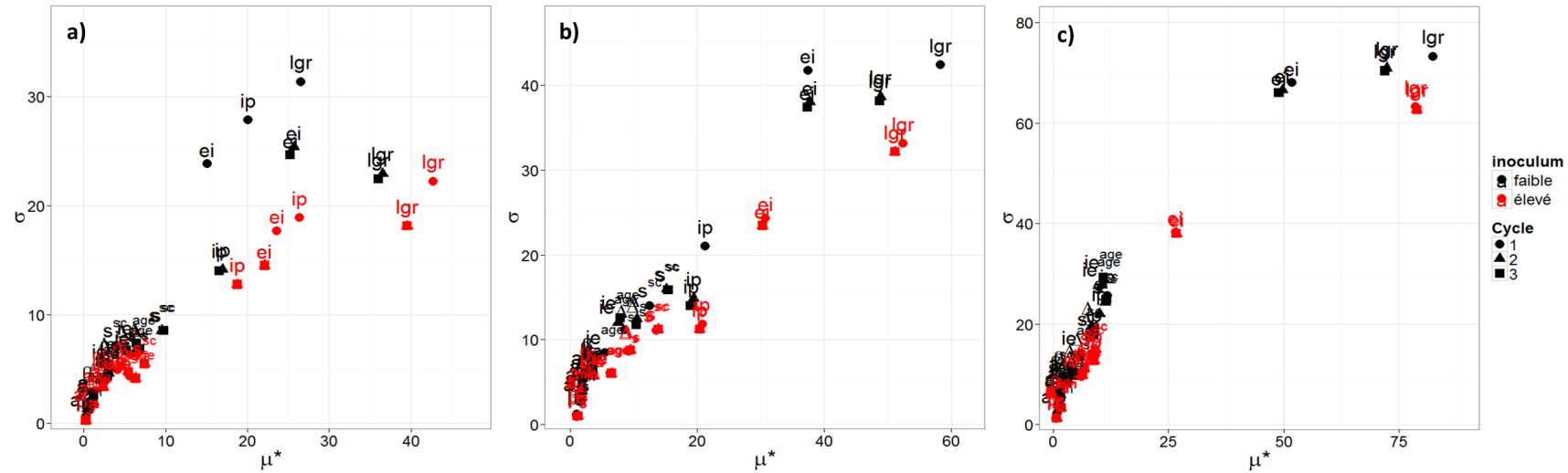


FIGURE A.11 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : pourcentage de maladie de la plante sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte.

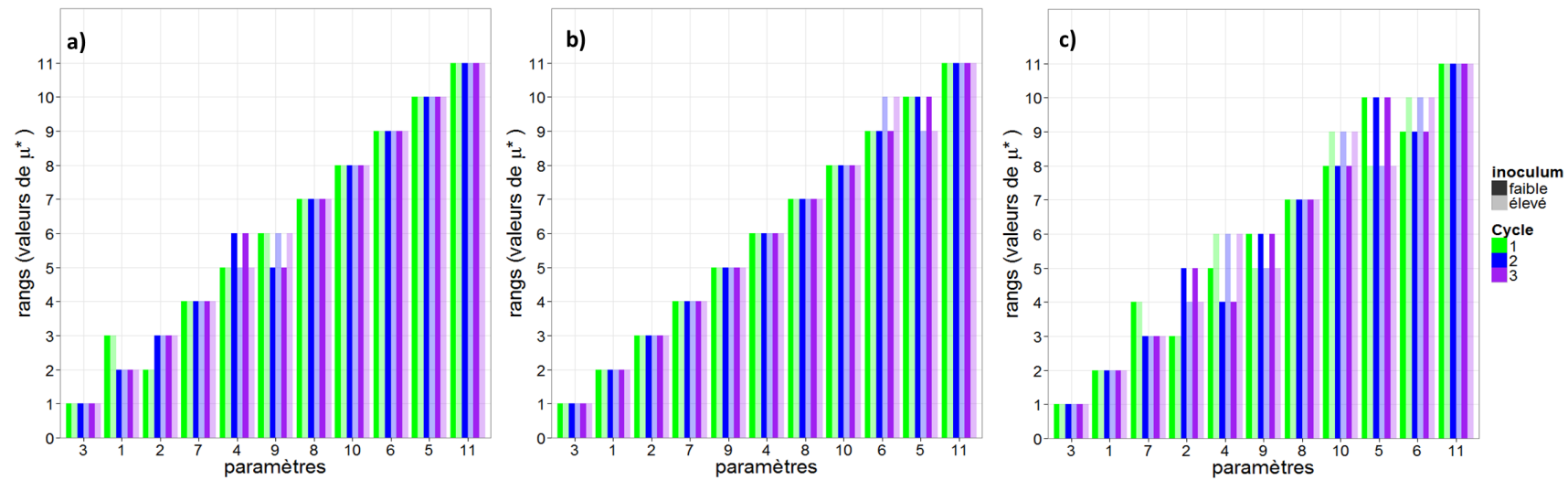


FIGURE A.12 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : pourcentage de maladie de la plante sur trois cycles de culture à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 11 désigne la valeur la plus faible. Numérotation des paramètres : 1: ei , 2: ip , 3: lgr , 4: ei^{age} , 5: a^{min} , 6: sc_a , 7: sc_s , 8: Δ_a , 9: Δ_s , 10: β_a , 11: β_s .

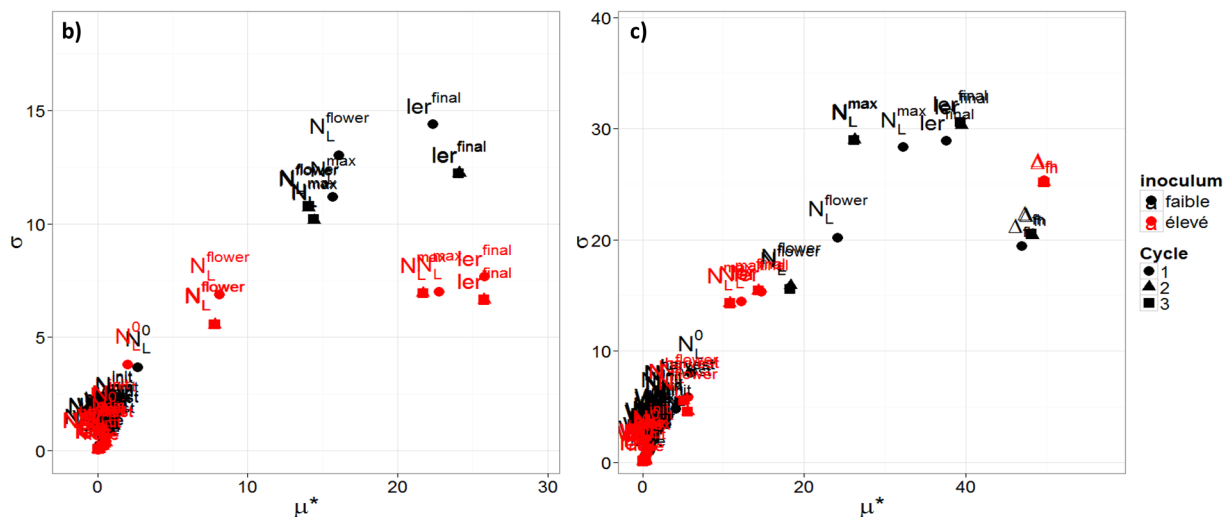


FIGURE A.13 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : sévérité de la plante sur trois cycles de culture à deux dates, b) floraison et c) 30 jours avant la récolte.

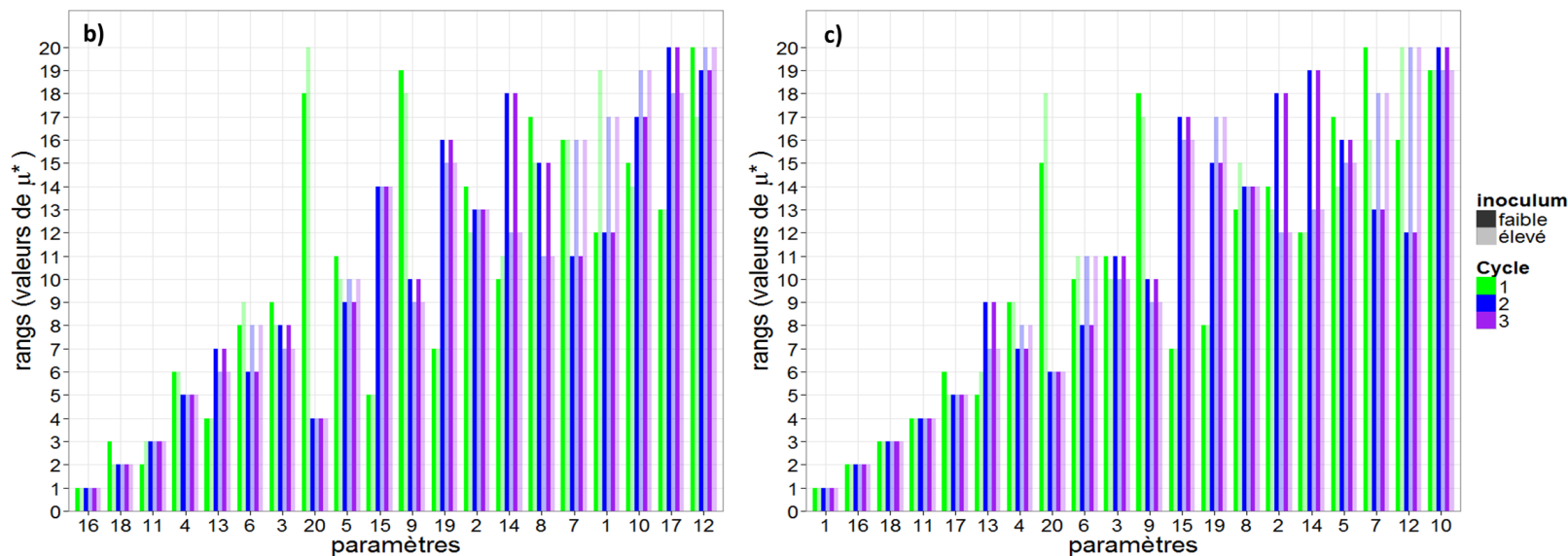


FIGURE A.14 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : pourcentage de nécroses sur la plante, sur trois cycles de culture à deux dates, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 20 désigne la valeur la plus faible. Numérotation des paramètres : 1 : Δ_{fh} , 2 : ϕ_L^V , 3 : L_L^{max} , 4 : L_L^{min} , 5 : W_L^{max} , 6 : W_L^{min} , 7 : ϕ_L^H , 8 : l_p^0 , 9 : l_p^1 , 10 : H_L^0 , 11 : N_L^{flower} , 12 : H_L^1 , 13 : N_L^0 , 14 : l_{node} , 15 : ler^{init} , 16 : ler^{final} , 17 : $N_L^{harvest}$, 18 : N_L^{max} , 19 : $N_L^{0,init}$, 20 : $N_L^{1,init}$.

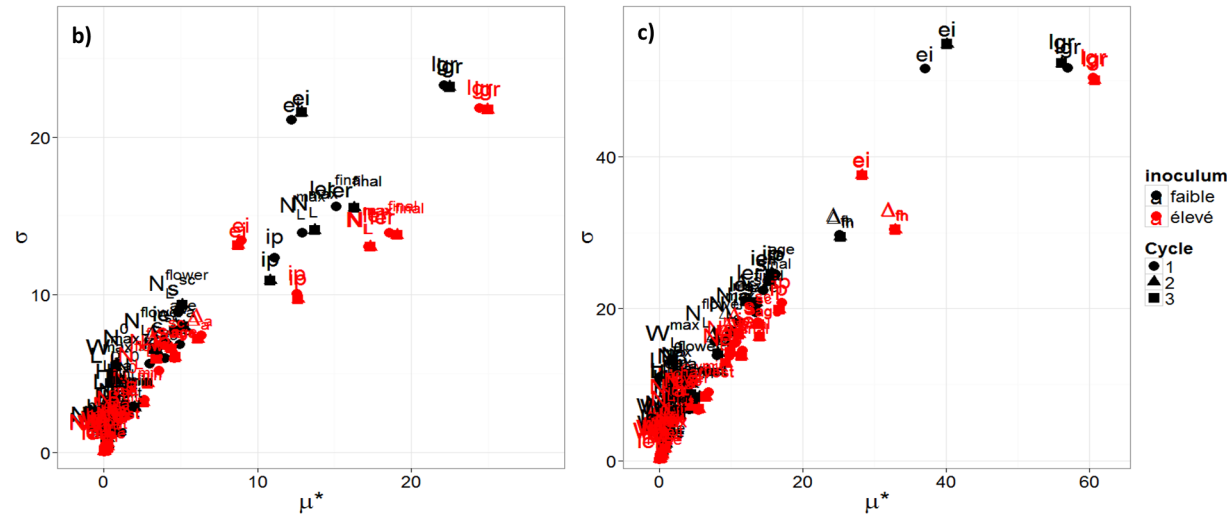


FIGURE A.15 – Écart-type des effets élémentaires des 11 paramètres du module épidémiologique en fonction de la moyenne des valeurs absolues des effets de ces paramètres sur la sortie d'intérêt : pourcentage de surface nécrosée sur la plante sur trois cycles de culture à deux dates, b) floraison et c) 30 jours avant la récolte.

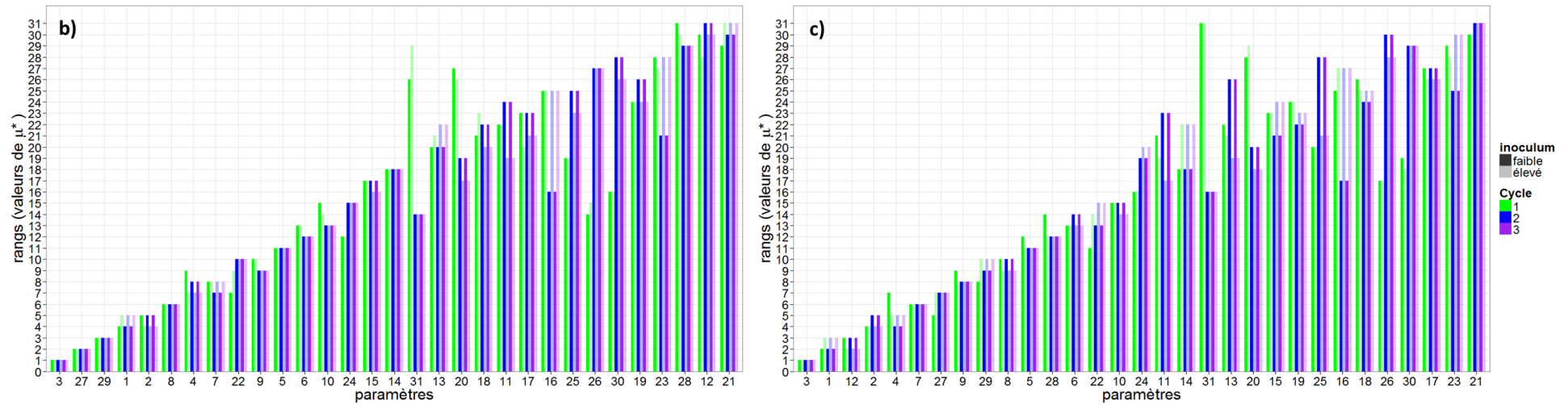


FIGURE A.16 – Classement des paramètres selon les valeurs de μ^* pour la sortie d'intérêt : pourcentage de nécroses sur la plante, sur trois cycles de culture à deux dates, b) floraison et c) 30 jours avant la récolte. Le rang 1 désigne le paramètre ayant la valeur de μ^* la plus élevée, le rang 20 désigne la valeur la plus faible. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s , 12 : Δ_{fh} , 13 : ϕ_L^V , 14 : L_L^{max} , 15 : L_L^{min} , 16 : W_L^{max} , 17 : W_L^{min} , 18 : ϕ_L^H , 19 : l_p^0 , 20 : l_p^1 , 21 : H_L^0 , 22 : N_L^{flower} , 23 : H_L^1 , 24 : N_L^0 , 25 : l_{node} , 26 : ler^{init} , 27 : ler^{final} , 28 : $N_L^{harvest}$, 29 : N_L^{max} , 30 : $N_L^{0,init}$, 31 : $N_L^{1,init}$.

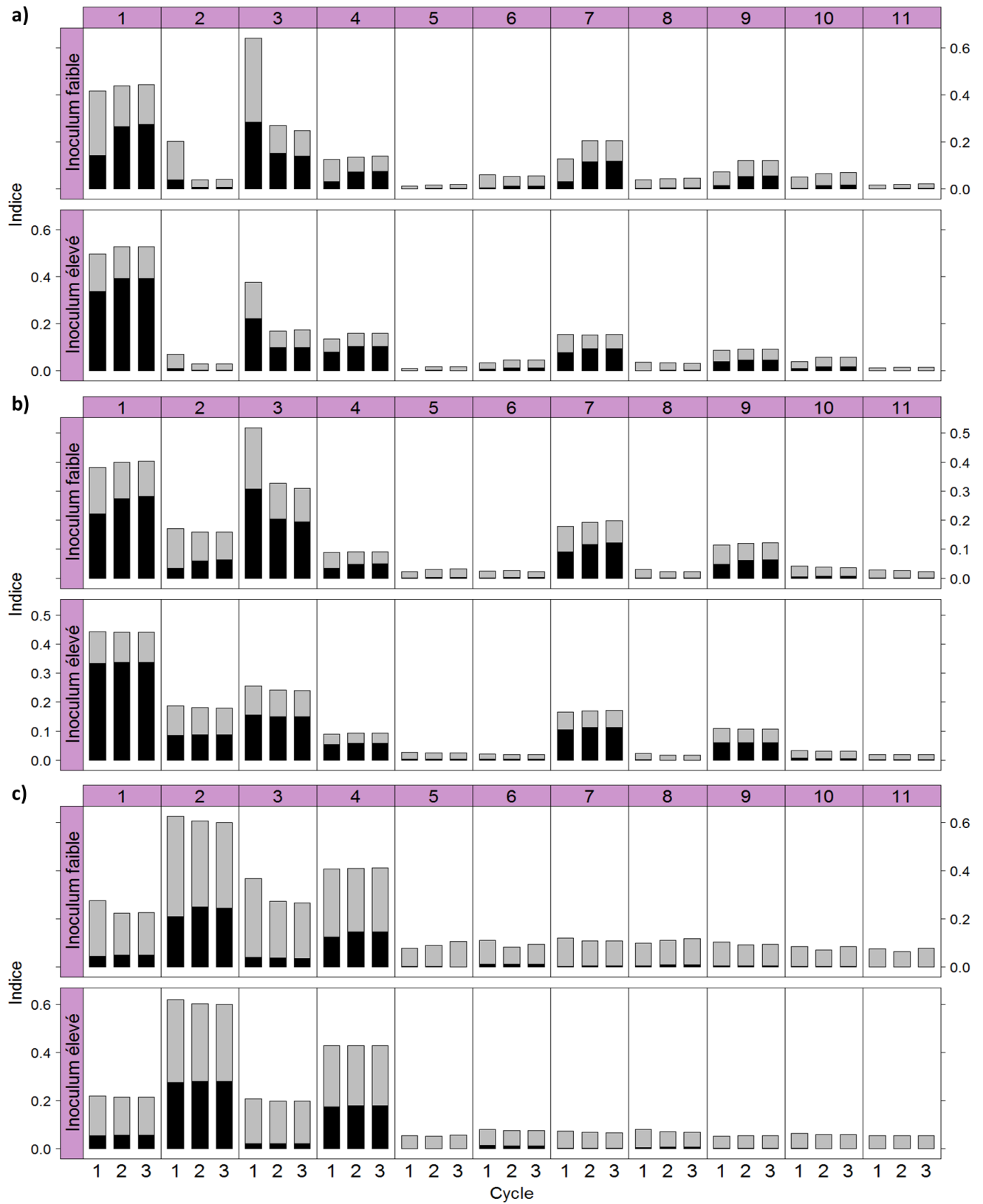


FIGURE A.17 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 11 paramètres du module épidémiologique calculés sur le nombre de lésions en incubation sur la plante à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s .

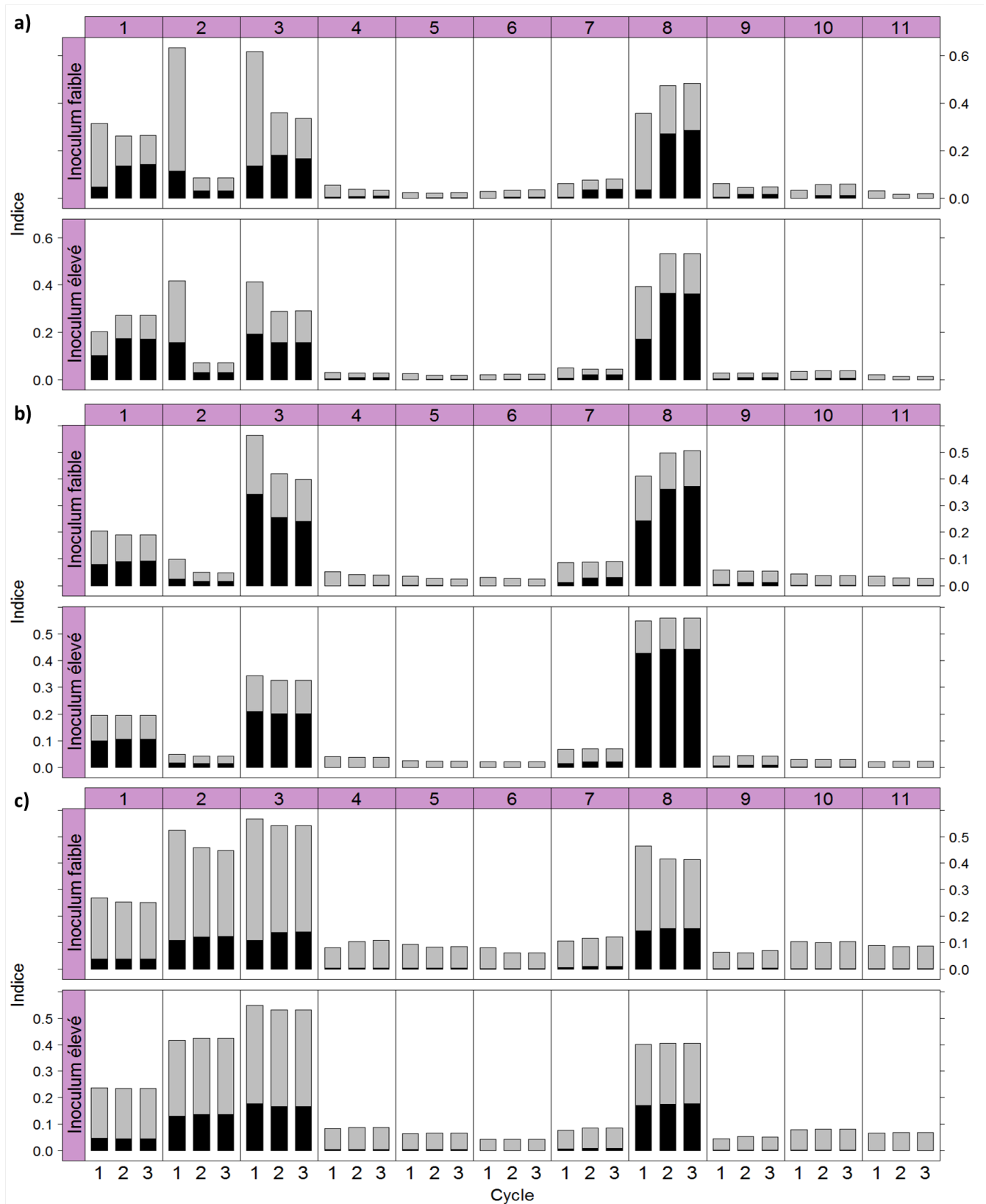


FIGURE A.18 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 11 paramètres du module épidémiologique calculés sur le pourcentage de surface conidienne sur la plante à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s .

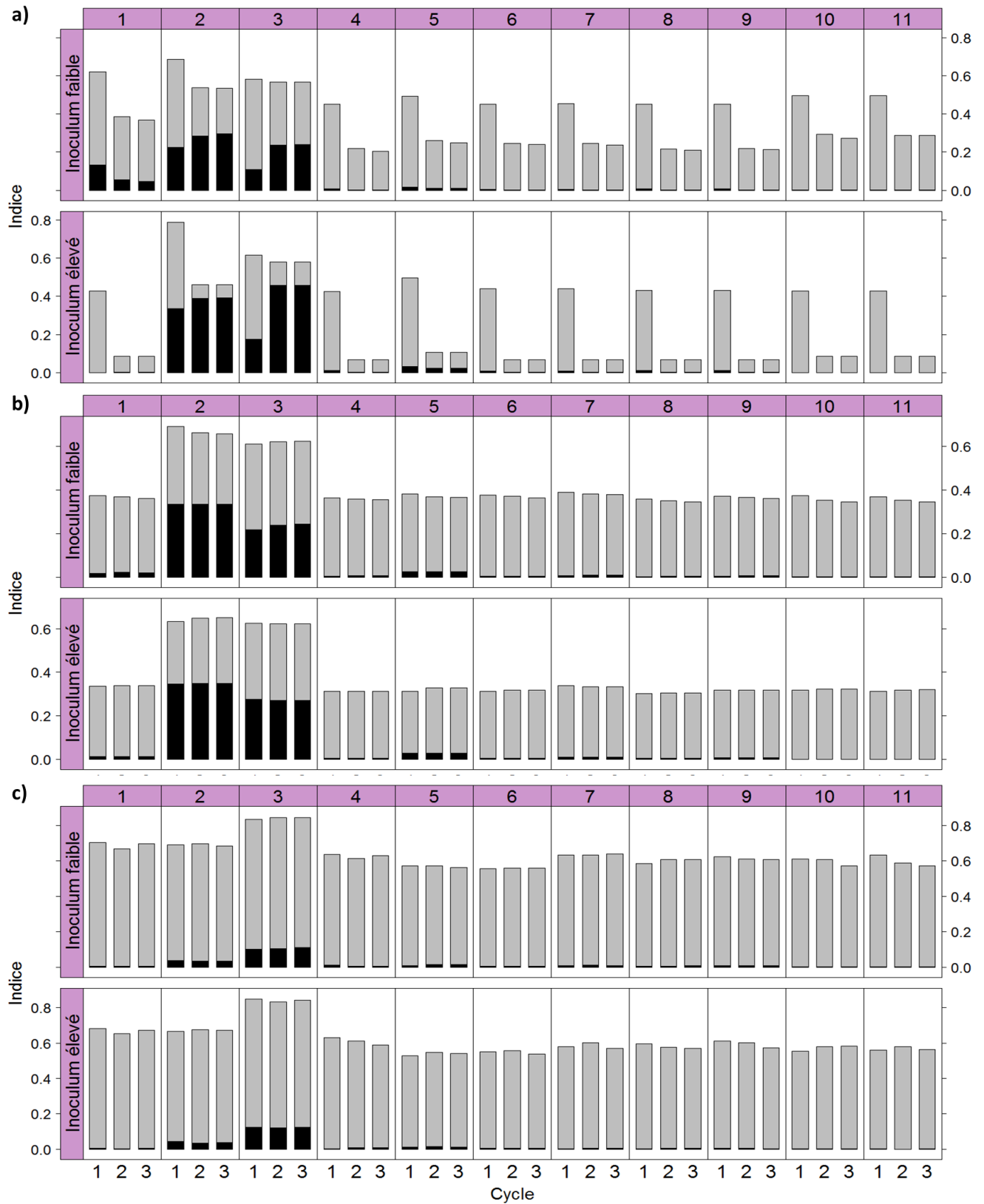


FIGURE A.19 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 11 paramètres du module épidémiologique calculés sur le rang de la PJFM sur la plante à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s .

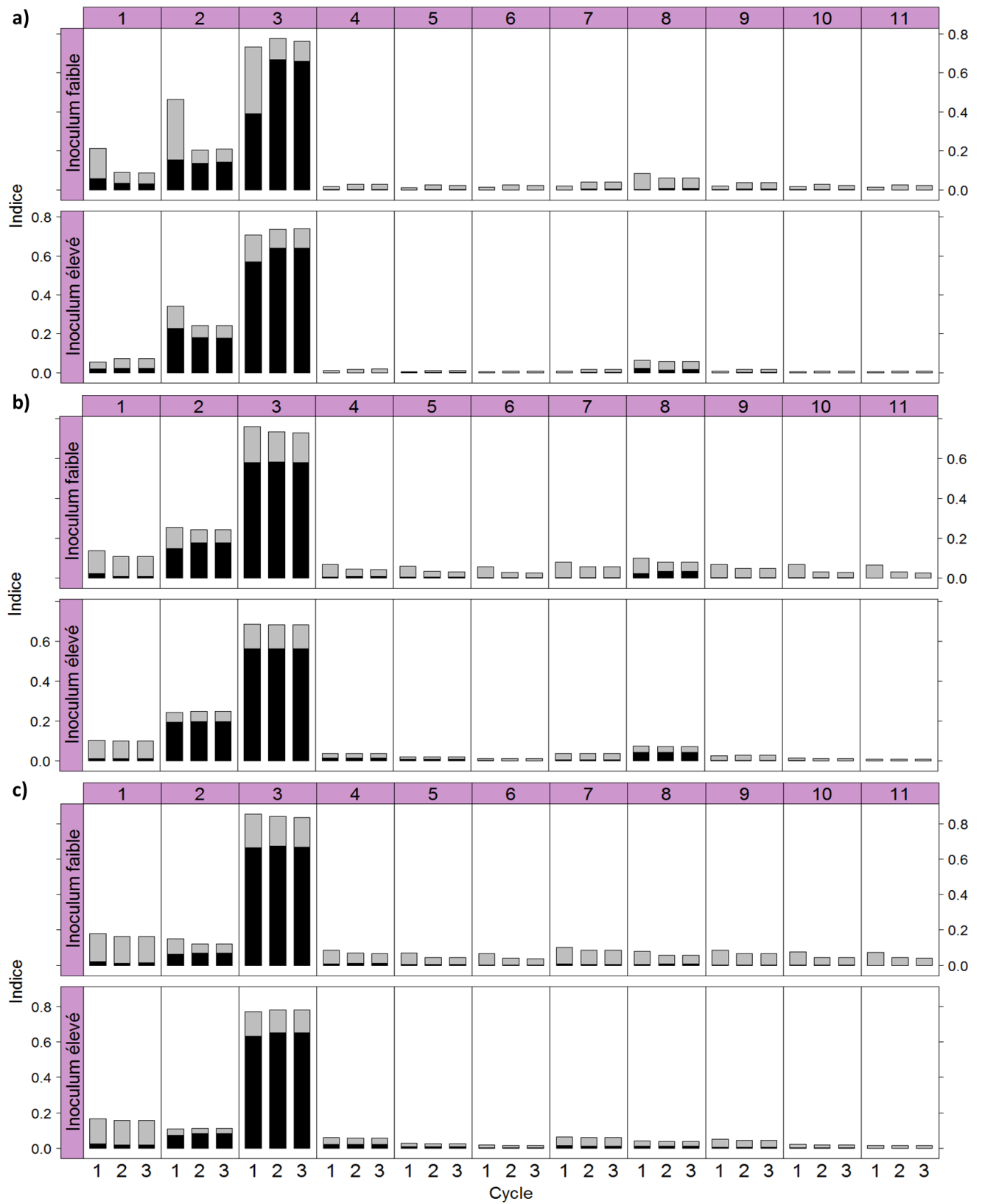


FIGURE A.20 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des 11 paramètres du module épidémiologique calculés sur le pourcentage de surface nécrosée sur la plante à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s .

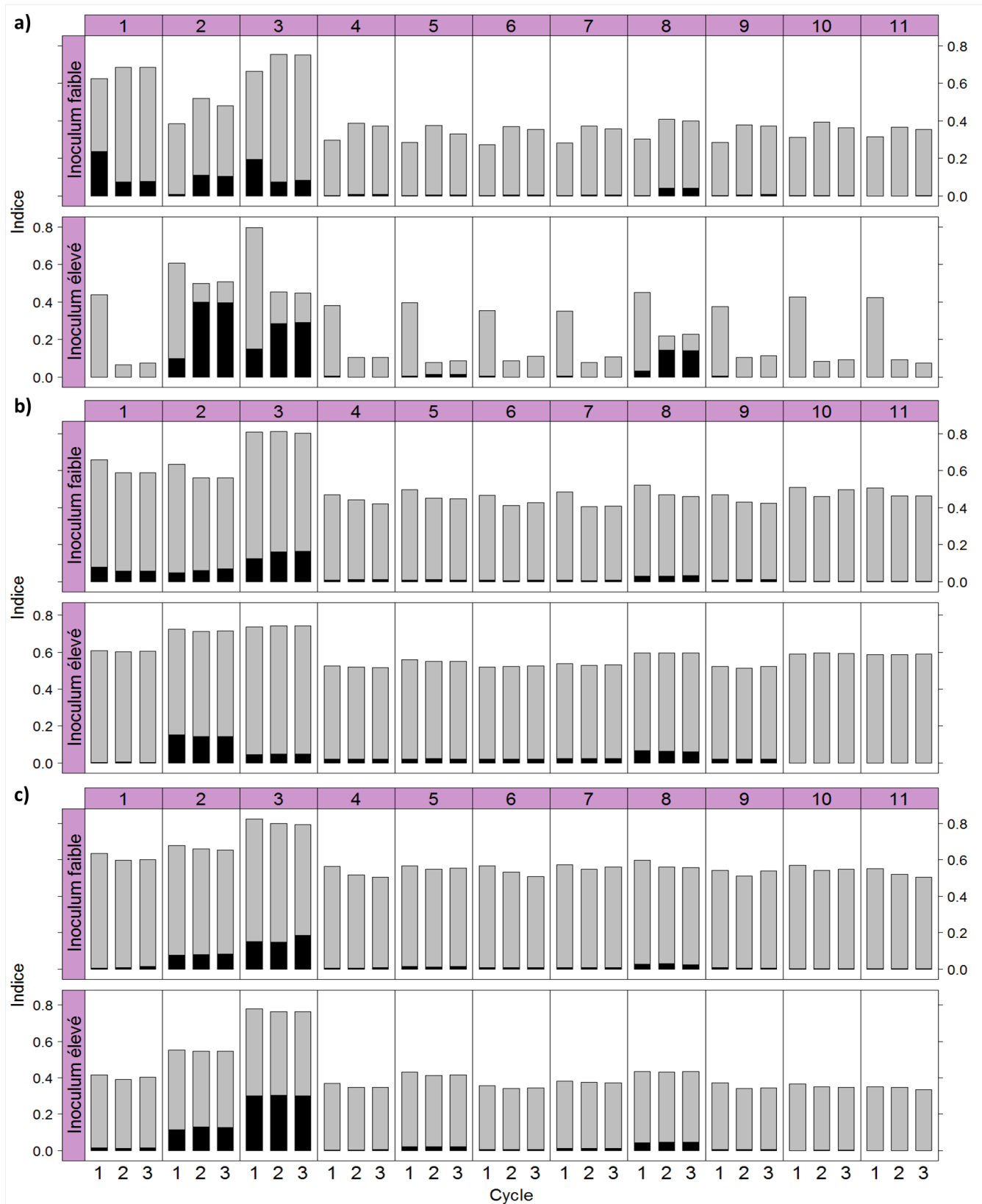


FIGURE A.21 – Indices de sensibilité principaux (en noir) et totaux (en gris + noir) des paramètres 11 paramètres du module épidémiologique calculés sur le rang de la PJFN de la plante à trois dates, a) initiation florale, b) floraison et c) 30 jours avant la récolte. Numérotation des paramètres : 1 : ei , 2 : ip , 3 : lgr , 4 : ei^{age} , 5 : a^{min} , 6 : sc_a , 7 : sc_s , 8 : Δ_a , 9 : Δ_s , 10 : β_a , 11 : β_s .

